Aalto University
School of Science
Degree Programme in Engineering Physics and Mathematics

Teemu Känsäkangas

Quality Assessment Methods for Traffic Incident Information

Master's thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Technology in the Degree Programme in Engineering Physics and Mathematics.

Helsinki, September 14, 2015

Supervisor: Professor Ahti Salo
Advisors: Tomi Laine M.Sc. (Tech.), Strafica Ltd.
Osmo Salomaa M.Sc. (Tech.), Strafica Ltd.

Aalto University
School of Science
Degree Programme in Engineering Physics and Mathematics

ABSTRACT OF MASTER'S THESIS

Author:	Teemu Känsäkangas		
Title:	Quality Assessment Methods for Traffic Incident Information		
Date:	September 14, 2015	**Pages:**	ix + 75
Major:	Systems and Operations Research	**Code:**	Mat-2
Supervisor:	Professor Ahti Salo		
Advisors:	Tomi Laine M.Sc. (Tech.) Osmo Salomaa M.Sc. (Tech.)		

Road incidents are a significant factor in traffic delays. In Finland, approximately 65 % of traffic delays are connected to different types of incidents. While incidents themselves are difficult to eliminate altogether, their impact can be diminished with precise incident reporting.

EIP (European ITS Platform) project, invoked by the ITS (Intelligent Transportation Systems) directive adopted by the European Parliament and the Council, states explicit quality criteria for different traffic information types. This thesis is focused on the reporting quality of traffic incidents. The criteria set for real-time traffic information (RTTI) includes spatial, temporal, and spatio-temporal measures. Incident reporting quality measures are implemented according to the defined criteria and tested at selected locations in the Finnish road network.

The traffic incident information quality assessment was performed by comparing incident data provided by the Finnish Transportation Agency to detected incidents from HERE traffic service and emergency mission data from the Finnish Emergency Response Centre. Based on the quality assessment results, this thesis provides the level of reporting quality of traffic incidents in Finland.

The developed incident detection methods proved very promising but require more study. Based on the observed detection performance, the speed data alone was deemed insufficient for incident detection baseline data. Despite the difficulties, criteria calculations could be performed and the overall incident information quality in Finland achieved basic levels. The basic principles in the incident detection process can also be used to detect traffic incidents in real-time, provided that additional traffic data can be obtained in the future.

Additionally, this thesis proposes a solution for incident information quality assurance which can be used to monitor the traffic incident information quality against the defined criteria quality levels. Control charts and a special four-field presentation methods were defined for traffic incident information monitoring. These tools provide a real-time view of the incident information quality alongside the defined quality levels and show typical criteria value configuration cases.

Keywords:	traffic, traffic incident, quality assessment, quality criteria
Language:	English

Aalto-yliopisto
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan koulutusohjelma

DIPLOMITYÖN TIIVISTELMÄ

Tekijä:	Teemu Känsäkangas		
Työn nimi:	Laadunarviointimenetelmät liikennehäiriötiedoissa		
Päiväys:	14. syyskuuta 2015	**Sivumäärä:**	ix + 75
Pääaine:	Systeemi- ja operaatiotutkimus	**Koodi:**	Mat-2
Valvoja:	Professori Ahti Salo		
Ohjaajat:	Diplomi-insinööri Tomi Laine Diplomi-insinööri Osmo Salomaa		

Liikenneonnettomuudet ovat merkittävä tekijä liikenneruuhkissa – noin 65 % ruuhkista johtuu erityyppisistä liikenneonnettomuuksista. Vaikka liikenneonnettomuuksia on hankala ennaltaehkäistä, niiden liikenteellistä vaikutusta voidaan vähentää tarkalla ja ajantasaisella häiriötiedottamisella.

Euroopan parlamentin omaksuman ITS-direktiivin synnyttämässä EIP-projektissa (European ITS Platform) on määritelty eri ajantaisen liikennetiedotustyyppien laatukriteeristöt. Tässä työssä tutkitaan ajantasaisen häiriötiedotuksen laatua. Häiriötiedotuksen (real-time traffic information) kriteeristöön kuuluu paikannuksellisia, ajallisia ja näitä yhdisteleviä laatutekijöitä. Työssä toteutetaan laatumittareiden laskuperiaatteet kriteeristön perusteella, ja niitä kokeillaan eri tieverkon osilla Suomessa.

Häiriötiedotuksen laadunarviointi toteutettiin vertaamalla Liikenneviraston tuottamia häiriötietoja HERE-karttapalvelusta tunnistettuihin häiriöihin ja Hätäkeskuslaitoksen hälytystehtäväaineistoihin. Saatujen tulosten pohjalta selvitettiin häiriötiedotuksen laatutaso Suomessa.

Työssä kehitetyt häiriöntunnistusmenetelmät osoittautuivat lupaaviksi, vaikka ne vaativat lisää kehittämistä. Häiriöntunnistuksen tarkkuudesta päätellen nopeusaineisto yksistään ei näytä riittävän häiriöntunnistuksen pohja-aineistoksi. Vaikeuksista huolimatta kriteerien arvot pystyttiin laskemaan, ja tuloksena Suomen häiriötiedottamiselle määriteltiin perustasoa oleva laatuluokka. Häiriöntunnistuksen perusperiaatteita voidaan myös hyödyntää täysin reaaliaikaisen tunnistuksen kehittämisessä, mikäli tunnistusta helpottavaa lisäaineistoa on saatavilla tulevaisuudessa.

Tässä työssä tutkittiin myös erilaisia häiriötiedon laadunvarmistukseen liittyviä menetelmiä, joilla pystytään seuraamaan häiriötiedotuksen laatua kriteeristön laatutasoihin nähden. Tuloksena syntyi kontrollikortteihin ja erityiseen nelikenttäkuvaan perustuvat laadunvalvonnan tarkastelutavat, joilla häiriötiedotuksen laadun kehitystä voidaan seurata laatutasoihin verrattuna, ja jotka mahdollistavat häiriötapausten eri kriteerikonfiguraatioiden tarkastelun.

Asiasanat:	liikenne, liikennehäiriö, laadunarviointi, laatukriteerit
Kieli:	Englanti

Acknowledgements

I wish to thank my supervisor, advisors, and the EIP+ project steering group who provided critical assistance for this thesis. Your comments and steering helped me to achieve something I could not manage on my own. Ahti, I knew from the start that I chose the right supervisor for my thesis. Tomi and Osmo, your combined transportation engineering and systems analysis expertise was simply the best nest of knowledge I could feast on. Risto, Aapo, Sami, Jani, and Anna, our steering group meetings were a crucial part not only for the success of the project involved but also for this thesis.

Years of studies would be too much for anyone without proper freetime activities. I want to and will thank my fellow Polytech Choir singers for keeping my spirits up on our Wednesday rehearsals. Our music-making and all the laughter kept me going. Paavo, Eero, Lastu, Allu, and Valtteri, Stagement Pro is something we should work on next year.

My family were very supportive with my studies all these years. Not even the great distance between me and my mother, father, brother, and sisters diminished the support and love got from them. Mutsi and faija, you were always very supportive of my studies all the way from the primary school. I think you are truly a perfect specimen for parents! Sisters Tiia and Tuija, your so-bad-it-is-actually-good humour gave me crippling laugh fits from which I am still recovering. You two carry on with your captivating and positive-additude-inducing craziness! It is you, Tero, who especially inspired me to push on with my thesis and achieve something worth to be proud of. Ylipaljon kiitoksia teille!

Last and foremost I also want to thank my girlfriend for her existence in my life. Maija, without you my life beyond studies would have been much emptier. You supported me on my darkest times. Your unconditional love and understanding was and is overwhelming. I love you.

Thank you all and let us continue to the next chapter.
Helsinki, September 14, 2015

Teemu Känsäkangas

Abbreviations and Acronyms

AlertC	A language-independent exchange format for traffic information in the RDS-TMC channel.
API	Application Program Interface.
ARIMA	Autoregressive Integrated Moving Average; a statistical model for estimating time-series data.
DATEX II	A standard for traffic information exchange between traffic information centres.
EIP	European ITS Platform; a project in which a quality criteria palette was defined for different traffic information types.
EIP+	EIP successor project in which previously defined quality criteria are tested in the EU member states.
FERCA	Finnish Emeregency Response Centre Administration.
FTA	Finnish Transport Agency.
ITS	Intelligent Transportation Systems.
QKZ	Spatio-temporal quality metric for general traffic information developed at BMW.
QRTTI	Spatio-temporal quality metric for real-time traffic information.
QSRTI	Spatio-temporal quality metric for safety-related traffic information developed at TNO, Netherlands.
RDS-TMC	Radio Data System - Traffic Message Channel; a technology for delivering traffic information to navigator systems.
RTTI	Real-time Traffic Information.
SRTI	Safety-related Traffic Information.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Road traffic is subject to continuous changes in demand in short and long timespans. Therefore, rigid road infrastructure alone will not adapt to fluctuating demands for road capacity. As a result, road network not only has to offer adequate capacity for the recurring traffic congestion but also flexibility for instance by providing alternative routes to avoid non-recurring congestions caused by traffic incidents. Road network could be scaled in a way that no recurring or even non-recurring congestions would ever occur but the scale would be difficult to estimate and the construction and maintenance would be very expensive. Thus, a typical, high-traffic road network is highly susceptible to congestions caused by traffic incidents.

Road incidents have a noticeable share in the typical causes of traffic delay – in Europe, the number is between 10 to 25 % and in Finland as high as 65 %. [31] As a result, a significant share of congestion is unavoidable without proper information about associated incidents. These non-recurring congestions can be avoided by implementing and utilising an efficient information platform in which congestion-causing incidents are presented as accurately as possible. This information is made available to public access point as well as to third-party service providers who present it in their navigation systems or other services. By using this processed information road users can avoid congestions, prevent additional incidents, and improve the overall efficiency of the traffic flow. As one access point example, Finnish Transport Agency (FTA) has all traffic incident content and other real-time traffic information visualised in a map-based web interface [7]. V-Traffic is a private company that has a similar web interface [17] as FTA but traffic information includes material from other sources, such as incidents logged by crowdsourcing.

The European Parliament and the Council have adopted a directive [18] which states several prioritised Intelligent Transport Systems (ITS) requirements. Among these requirements are definitions of reliable traffic information services and related actions to standardise these services across EU countries. European ITS Platform (EIP) project addresses certain parts of this directive by standardising traffic information quality control principles by defining quality criteria, levels of quality and monitoring requirements for participating member states. EIP+ project is a successor to the EIP project where quality criteria defined before is assessed and quality assessment methods are defined based on these criteria. Participating states study criteria on different information types. One of the incident types referred to as real-time traffic information (RTTI) is studied in Finland. Studies shown in this thesis are a part of this research led by the Finnish Transport Agency (FTA).

1.1.1 Incident Information Value Chain in Finland

Incident recording, reporting, and publishing process is an information value chain. In this context, the value chain characterises the general sequence of operations to get information from the traffic incident scene to the map view in navigators or web services. The value chain is divided into two main segments and associated subsegments that represent different phases in information flow, as can be seen in figure 1.1. Content segment includes detection and processing subsegments, in which states in traffic are transformed into traffic information by traffic authorities. Service segment includes provision and presentation, where traffic information from different access points are collected, manipulated, and visualised for end users.

In Finland, content segment is mainly governed and operated by the FTA and Finnish Emergency Response Centre Administration (FERCA). Traffic incidents are detected when an accident bystander makes an emergency call to the FERCA. After the dispatch of the rescue department and the police, the incident is automatically appended to an incident impulse queue in the FTA systems. In the FTA, operators empty the queue one item at a time by first providing preliminary information about the traffic incident in maximum 30 minutes after the arrival in the queue. After the rescue crew or police provides more information about the incident at the scene by phone, the operator fills additional details for the corresponding incident datum.

In the service segment of the value chain, the data provided is first collected from the access point, supplemented with additional information and adapted to fit the service in question. Both the preliminary and updated reports are sent to a separate Digitraffic service where the information is available through an interface. Also, the information is further sent from Digitraffic to web services, email subscribers, and RDS/TMC operators. After the data collection and manipulation, information is presented to end user via display on a car navigator, web interface, or other service. The process of the incident detection and information provision is shown in figure 1.2.

1.2 Objectives

1.2.1 Quality Assessment Methods

The main objective of this thesis is to assess the quality of the traffic incident information with a set of quality assessment methods and with data provided by the FTA and third-party service providers. Another objective is to study the qual-

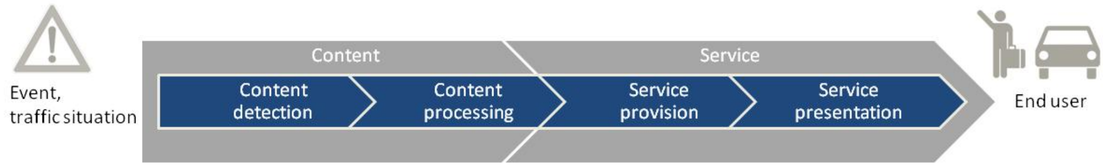


Figure 1.1: Real-time traffic information value chain. [22]

ity assessment criteria provided by the EIP project and provide recommendations for criteria development. The focus is in the content segment of the information value chain. In the content detection, the coverage and response time of the incident reports are studied. In content processing, the quality of the outcome based on the reports is assessed. Reporting quality in the service segment and interdependencies between segments are also briefly discussed.

The used quality assessment methods are based on the quality criteria proposed in the EIP project, listed in table 1.1. Timeliness and latency are temporal criteria which are used to assess the delays in the incident information value chain. Location accuracy is a spatial criteria for measuring the preciseness of the reported incident location. Error rate and event coverage are set criteria defined for the assessment of redundancy and completeness of the incident information. Potential methods for the quality measurement calculation and other related tasks are collected from the literature. The chosen methods are defined and modified to be compatible with Finnish traffic datasets.

1.2.2 Feasibility of Quality Assessment Criteria in Finland

Real-time traffic incident information quality criteria proposed in the EIP project are definitions of the quality principles that have to be used in the quality assessment. Concrete methods are defined according to these criteria. These methods are required to be implemented according to data available in the relevant country. Relevant quality criteria in incident reporting are listed in table 1.1.

The implementing organisation must take the available data into account. Following the definition of the quality assessment methods, this thesis analyses and discusses the feasibility of the corresponding criteria for Finnish traffic environment. The analysis also provides suggestions for corrections and supplements for the quality criteria.

1.2.3 Level of Incident Information Quality in Finland

The level of incident report quality in Finland over different criteria is defined by the quality levels presented in table 1.2. The locations are chosen on the principle that they accurately represent the diversity of regions and functional road types in Finland. Reporting quality analysis can only be performed on roads that produce sufficient amounts of traffic data. Therefore, the chosen set of roads consists only

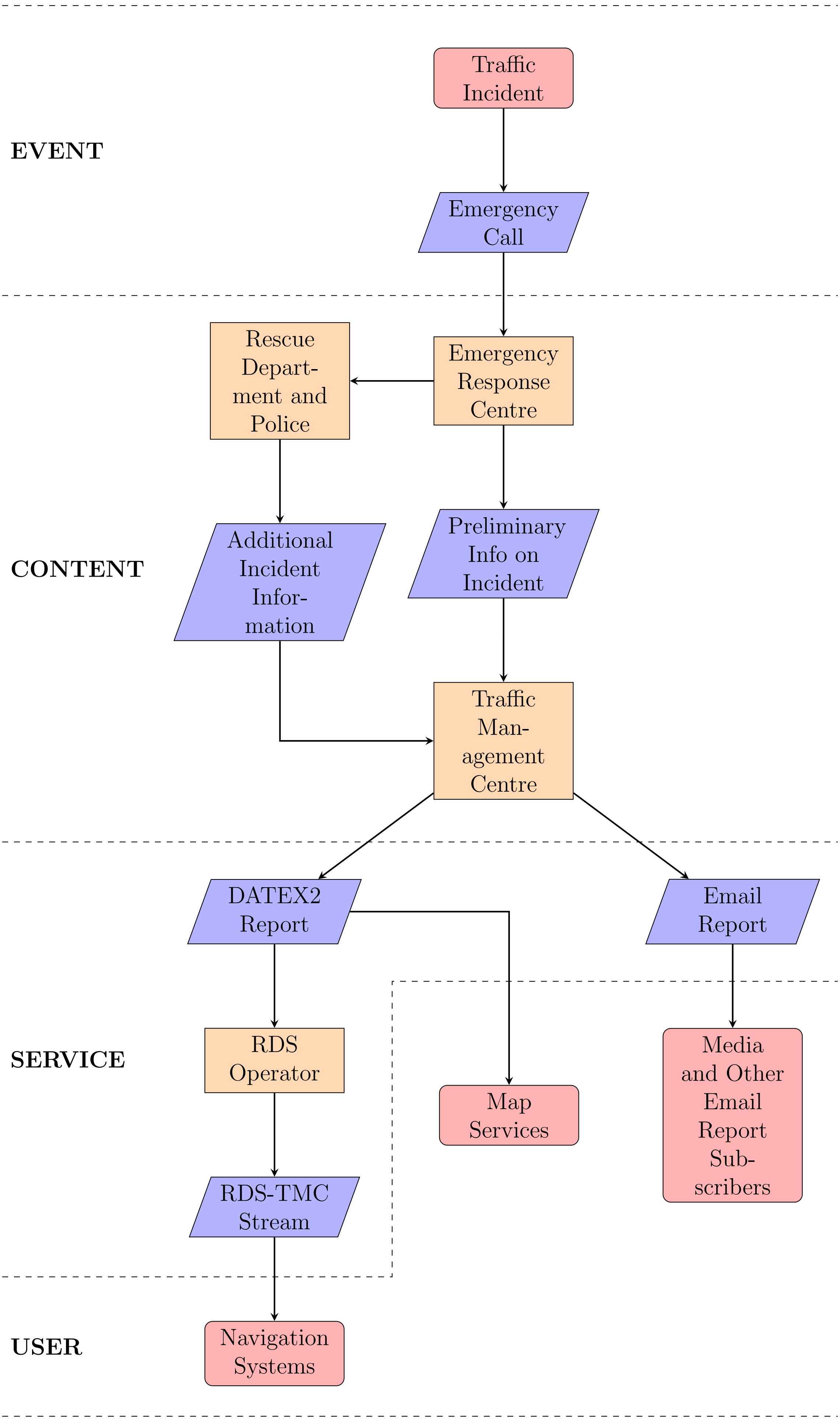


Figure 1.2: Traffic incident reporting process in Finland.

Table 1.1: Real-time traffic information quality criteria defined by EIP. [22]

Criteria	**Description**
Timeliness (start)	Delay between the occurrence of an event and the first (validated if necessary) detection of the event.
Timeliness (update)	Delay between the end or safety-relevant change of condition and the detection of this end or change.
Latency (content side)	The delay between the first (validated if necessary) detection of the event and the moment the information is provided by the content access point.
Location accuracy	The relative precision of the referenced location for the published event with respect to the actual location of the actual event.
Error rate	Percentage of the published events which are known to be not correct (concerning actual occurrence of this event type / class at the reported location at the reported time).
Event coverage	Percentage of the actually occurring events which are known to be correctly detected and published by type / class, time and location (i.e. Detection Rate).

of parts of the Finnish road network that involve at least moderate amounts of traffic volume and tendency to traffic incidents.

1.2.4 Definition of Quality Assurance Process in Finland

In addition to analysing the incident information quality, this thesis defines a quality assurance process that is suitable for the long-term monitoring of the incident information quality. The objective for the assurance process is to function as a tool for operators to monitor reporting quality in terms of satisfying the targeted level of the quality. Additionally, the assurance process is required to provide possible explanations for the problems in the incident information quality by providing information about the quality of individual criteria.

1.2.5 Context of the Study

The baseline data for the research in this thesis is collected from incidents occured in Finnish road network between October 15th, 2014 and February 28th, 2015 at 6 AM–9 PM. Additional data from other sources are collected from the same time span.

The studied incidents are restricted to motorways with a sufficient amount of daily traffic. The chosen road sections are mainly part of national highways (road numbers 1–12) and second level main roads (road numbers 45–170).

Incident types are filtered to only contain road accidents. For example construction sites, public events and other incidents known in advance are not studied.

Real-time traffic information consists not only of incident records, but also information about the state of the traffic, such as speed, congestion, occupancy, traffic incidents, and road works. Among these types, only real-time traffic incident information is studied.

Incident set to be analysed must be situated on a road network delimited to sufficiently representing set of locations, functional classes, average occupancy levels and levels of incident susceptibilities. Main roads and highways in Finland have been divided in operating environment classes that represent different functional classes, occupancy rates and incident susceptibilities. By selecting a sufficiently diverse set of operating environment classes from network, not only mentioned factors are covered but also geographical diversity is obtained. Operating environment classes in road network and associated explanations can be found in figure 1.3 and table 1.3. After studying the amount of incidents and traffic reactivity for incidents in different operating environment groups, classes P1, T2, and T4 are selected for quality assessment testing in this thesis.

Table 1.2: Level of quality parameters for real-time traffic information assessment criteria. [22]

Parameter	⋆ **(Basic)**	⋆⋆ **(Enhanced)**	⋆⋆⋆ **(Advanced)**	⋆⋆⋆
Timeliness (start) (95 %)	Best effort	Validation (if necessary) after first detection < 5 min	Detection & validation < 5 min after event occurrence	Detection & validation < 3 min after event occurrence
Timeliness (update/end) (95 %)	Best effort	Best effort	Detection & validation < 10 min after event change/end	Detection & validation < 5 min after event occurrence
Latency (content side) (95 %)	< 10 min	< 5 min	< 3 min	< 2 min
Location accuracy (area) (95 %)	Administrative region	Geographic area; 10 km accuracy	Geographic area; 5 km accuracy	Geographic area; 2 km accuracy
Location accuracy (road) (95 %)	Link between intersections	Link between intersections	< 5 km	< 2 km
Error rate	< 15 %	< 10 %	< 10 %	< 2 %
Event coverage	Best effort	> 90 % of all detected (and valitaded if necessary) events	> 80 % of all occuring events	> 90 % of all occuring events

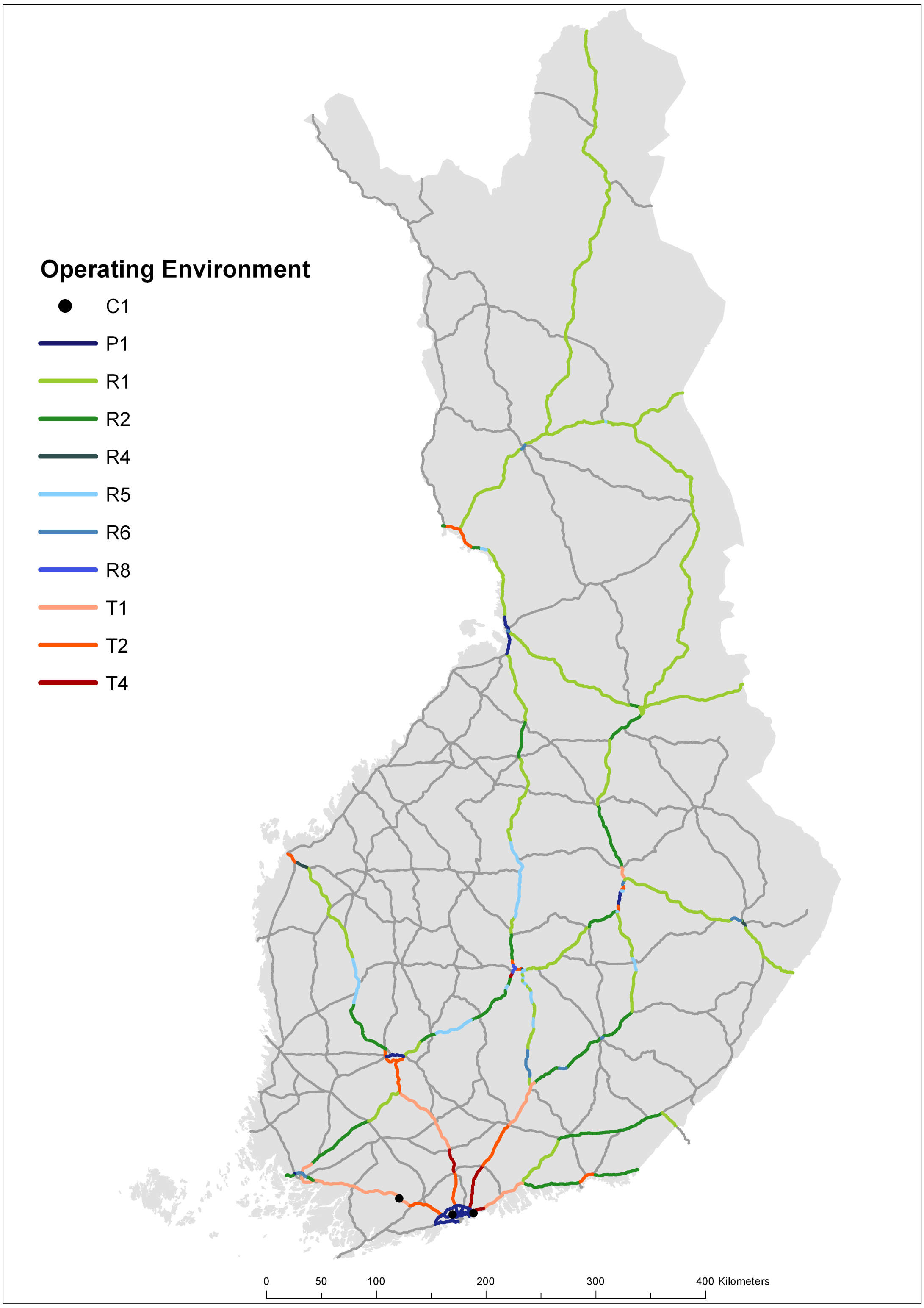


Figure 1.3: Finnish road network operating environment classes.

Table 1.3: Operating environment class glossary.

OE class	**Description**	**Flow-related problems**	**Safety-related problems**
P1	Peri-urban motorway or road interfacing urban environment		
T1	Motorway (link)	No	No
T2	Motorway (link)	No	Yes
T4	Motorway (link)	Yes	Yes

Chapter 2

Data and Tools

2.1 Incident Information

2.1.1 Finnish Transport Agency

Incident records from Finnish Transport Agency Digitraffic service are listed as a table-like structure where every row represents a single incident. Incident data to be assessed is made available for public use via interface [3] by Infotripla.

FTA uses HäTi (Häiriötietojärjestelmä; in English: Disruption Information System) to log traffic incidents and other RTTI event type information, such as road works notices. [6] HäTi system main screen has a queue where new incident impulses are listed. The system is operated by a team of operators who empty the impulse queue by filling the preliminary announcements of the incidents using a dialogue. Afterwards, more information is gained from authorities at the incident scene and it is filled and announced, or, if no additional information is available, the incident is either too short-term to be reported or considered a false alarm.

When the actual incident information is filled, the operator also provides more accurate details about the location of the incident. This is done in the map view. The operator is given the approximate location of the incident, between two TMC points on road network. Based on information from authorities on scene, the operator defines a more accurate location for the incident. An incident log time stamp is recorded when the information dialogue is opened.

When first announcements or updates to a previously occured incident are registered in a HäTi system, they are automatically sent to Infotripla interface for public use. The delay in submission to the interface is mainly dependant on the impulse queue processing speed of the HäTi operator. The order of magnitude for delay is typically few minutes.

All traffic information operated by in the HäTi system is sent to a separate Digitraffic service that provides an interface for accessing the data. The data is structured in a DATEX II scheme [2] which is a standard exchange format for traffic information in EIP member states. The data is provided as a XML file in which different data fields can be utilised to provide different details of the state of the traffic. The typical structure of a DATEX II file can be seen in appendix A. The location and the description of the traffic incident are referenced with AlertC formatted data. The AlertC is a format which is used in TMC radio channels and

Table 2.1: Location reference types. [27]

Name	Description
AlertCMethod2Point	A TMC node
AlertCMethod2Linear	A road section between two TMC nodes
AlertCMethod4Point	A point in a road section from a primary TMC node with a offset towards a secondary TMC node
AlertCMethod4Linear	A road subsection between a primary TMC node with a offset towards a secondary TMC node
AlertCMethodArea	An administrative area by munincipality or country code

is designed to be low in both bandwidth and storage space usage. In Finland a typical AlertC message is in form:

```
Tie 5, Kuopio. ENSITIEDOTE LIIKENNEONNETTOMUUDESTA.Tie 5 välillä
Mikkeli - Kuopio.Paikka: Kuopio. Noin 7 km ennen paikkaa Kuopio.
Tarkempi paikka: Jynkän liittymä. ENSITIEDOTE LIIKENNEONNETTOMUUDESTA.
Kesto: 19.10.2014 23:18http://liikennetilanne.liikennevirasto.fi
Liikenneviraston tieliikennekeskus TamperePuh 020*******Faksi 020*******
Sähköposti **********@liikennevirasto.fi
```

The translation is as follows:

```
Road 5, Kuopio. PRELIMINARY TRAFFIC INCIDENT REPORT.Road 5 between
Mikkeli - Kuopio.Location: Kuopio. Approx. 7 km before location Kuopio.
Precise location: Jynkä junction. PRELIMINARY TRAFFIC INCIDENT REPORT.
Duration: 19.10.2014 23:18http://liikennetilanne.liikennevirasto.fi
Finnish Transportation Agency road traffic department Tampere
Tel 020*******Fax 020*******Email **********@liikennevirasto.fi
```

Basically all necessary information about the incident is listed on the description, such as time, place, type of the incident and contact information. The location of the incident is indicated with five possible methods, which are described in table 2.1. This thesis studies all non-area location references which can be studied on a road section.

2.1.2 Finnish Emergency Response Centre Administration

Finnish Emergency Response Centre Administration is responsible for emergency mission dispatch in Finland. Emergency mission history records are kept on administration databases. There is a possibility to request a list of emergency missions in Finland for scientific purposes. Data is collected by Emergency Response Centre [4] representatives according to submitted specifications for desired time interval and location. The dataset includes the location of the incident and timestamps of the emergency call and the emergency personnel dispatch.

2.1.3 V-Traffic Twitter Feed

Twitter is used to gather data from V-Traffic, which is a third-party traffic service providing congestion and incident information for consumers and businesses. [17] Incident data is collected from V-Traffic twitter feed, where the time, and place of the incident is listed along with Twitter-specific additional data such as retweet and favourite counts.

2.2 Traffic Speed Data

Traffic flow describes the state of traffic in different network sections at given moments in time. Descriptive values vary depending on the source, but they usually include information about speed and congestion on network segments. Flow data can be connected to road network geometry with unique identifiers which are found in both mutually compatible sources.

2.2.1 Nokia HERE

Traffic speed data from Nokia HERE [5] is used for incident ground truth data detection reviewed in section 3.2.1. In HERE speed data, actual and free flow speeds are provided as well as congestion labelled as 'jam factor'. Additionally, a confidence factor describing the amount of samples is included for statistical purposes. Confidence factor defines the source for the measurements according to its value. Source for flow values is defined as

i. $CN > 0.7$: actual measured values,

ii. $0.5 < CN \leq 0.7$: values estimated from historical data, or

iii. $0 \leq CN \leq 0.5$: speed limits.

Traffic data is asynchronously updated on HERE traffic network links in three-minute intervals, which is adequate for data analysis in this thesis. The data has approximately 1.5–3 minute delay in relation to the real world state, which must be taken into account when performing analysis.

2.3 Road Network Geometry

In order to analyse traffic flow variations in specific road sections and to visualise traffic flows in a graph, road geometry is needed. This geometry can be presented in two different forms: lines or segments. Lines are structured as a series of coordinates that govern the line path. This way, various measurement sections that incorporate curves and turns can be categorised in a single identifier. Segments are presented as a list of start and end coordinates and thus can be studied as a single parts of a given line.

Static information in road network data contains measurement unique segment identifiers, and the direction of the traffic in segment which can be used to connect traffic flow data to said network. Static information varies between different

Table 2.2: Hardware specifications.

Component	Specification
Operating system	Debian 7.8 GNU/Linux
Processor	64-bit Intel®Xeon®E5-4620 2.20 GHz, 64 cores
Memory	128 GB 1333 MHz DDR3
Hard drive	225 GB solid state drive, 1.8 TB physical network hard disk drive

road network sources which makes cross-compatibility practically non-existent. Therefore, static information and actual, varying data always have to be used in conjunction.

2.4 Software Tools and Hardware

R is an open-source statistical programming language that is widely used for statistics, simulation, model estimation and other mathematical tasks, as well as data manipulation and visualisation. R can be extended in functionality with various packages which are available for wide range of applications.

In the ground work for this thesis, R is used for all necessary calculations and data manipulations. Several additional packages for R are installed to provide necessary functionality for methodology testing and data manipulation. [1] [9] [10] [12] [13] [14] [15]

The calculations and algorithms are run with a powerful 64-bit, 64-core server with Debian 7.8 Wheezy GNU/Linux operating system. Full system specifications are listed in table 2.2.

2.5 Data Collection

Data is collected from various sources. Basically, two types of data are used for assessment testing: traffic incident data and traffic speed data. The former will be used directly to assess EIP quality criteria, the latter will be used to detect incidents, which could not otherwise be obtained from other sources. The question whether to use direct incident data, or incidents detected from the speed data – or both, will depend on how successful detections are with studied detection algorithms.

Traffic incident data to be analysed will be downloaded from Infotripla interface, which provides the data in XML format. This data is converted to SQLite database format table where one row represents one incident and associated XML tags. Data can be read from the table by standard SQL commands.

FTA traffic incident data from Infotripla is missing information about the moment of time when the incident information arrives at the HäTi system incident impulse queue. Approximation for this value is the moment when the Emergency Response Center receives the distress call from the incident scene and the first release of the incident is automatically submitted to the HäTi servers. This information is collected from the Finnish Emergency Response Center emergency mission data and combined with the FTA data.

HERE traffic speed data is collected from HERE API, in XML format. This data is also converted to SQLite table where measurements in every link in both directions in HERE network at every time stamp in the study period is divided into rows in the table. This table can be accessed with standard SQL commands. Rows in these tables are numerous, and indexing typical to SQL languages enables to sample data with greater speed than parsing XML data.

V-Traffic incident records are collected from the Twitter feed first described in section 2.1.3. Most of the data in the feed is aggregated from FTA access point so it will mainly be used for supportive tasks. This feed is saved as an R data frame, and can be loaded later for analysis.

The choice for ground truth data will be determined after pre-testing of methods. The null hypothesis is that HERE traffic speed data and Emergency Response Center data could be jointly used as a ground truth.

Chapter 3

Literature

3.1 Previous Studies in Traffic Information Quality Assessment

Texas Transportation Institute conducted a research for the best practices for assessing data quality. Three types of user contexts were studied: real-time traffic data collection and publishing, historical traffic data collection and monitoring, and other industries such as management information systems and geospatial data sharing. The study concluded that the recommended quality criteria for measuring traffic data quality are [32]

- *Accuracy.* The amount of data values which are correct in relation to the source.
- *Completeness.* The amount of data that is present in the required attribute fields.
- *Validity.* The proportion of data which satisfy the requirements for the associated attribute fields.
- *Timeliness.* The amount of data that is provided at the required time.
- *Coverage.* The amount of data values in a sample that accurately depict source values.
- *Accessibility.* The relative ease of data retrieval and manipulation.

Three of the listed requirements can be connected to the criteria palette defined in the EIP. Accuracy is closely related to the location accuracy criterion while timeliness has equivalent metric with the same name in EIP criteria. The coverage requirement can be connected to both error rate and event coverage criteria. The remaining three requirements – completeness, validity and accessibility – are present for data structural quality, readability, and conformity to the traffic reporting guidelines. These additional metrics should be considered when revising the criteria palette in the EIP+ successor projects.

3.2 Ground Truth Alternatives

One prequisite for quality assessment is to have a comprehensive set of data for comparative analysis. This data – referred to as 'ground truth' – forms a representation of the assumed real state of traffic against which the analysed data is tested.

The most important requirement for ground truth is that it represents the real-world traffic as accurately as possible. That is, ground truth incident occurring time and location should be as close to actual, realised values as possible. Another desired feature for ground truth is that it should be independent from the sources that are assessed.

3.2.1 Incident Detection from Traffic Data

One way to collect traffic incident ground truth is to detect incidents from real-time traffic status data. This approach ensures that the detected incidents not only represent the most accurate moment in time when the incident has had effect on the traffic flow, but also that the ground truth is not dependent of any other incident reporting systems or human factors.

Traffic data collection approaches can be divided into two main categories, fixed and mobile. Measuring devices integrated into the road infrastructure are used in the fixed data collection, such as detector loops and traffic cameras. In mobile data collection, the state of the traffic is collected from road users' cellphone navigation applications, navigators or other systems capable of measuring and sending location and speed to respective service provider systems.

In Finland, fixed data traffic state data collection of the traffic and speed in a cross section can be measured with LAM points (liikenteen automaattinen mittausjärjestelmä; in English: automatic measuring of traffic) [8] which are induction detector loops recessed into the road pavement. Traffic occupancy information could be very helpful addition to traffic speed data in incident detection because these two measures complement each other in calculating the impact of the traffic incident in upstream road sections. However, LAM point coverage in Finnish road network is very sparse especially in non-urban areas. Therefore, incident detection based on LAM measurements is not a viable choice.

There are several providers for mobile floating vehicle data and other types of real-time traffic information in the market, such as HERE, TomTom, and Google. The quality of the floating traffic data is based on the amount of service users, the granularity of the measuring segments in the road network, and the robustness of the service providing infrastructure. One favourable aspect of the floating data is that the associated services are essentially international and provide same source of ground truth for every country wanting to utilise methodology presented in this thesis. However, floating traffic information services usually measure only traffic speeds. This limits the palette of possible incident detection methods that can be used.

In this thesis – following the study of the possible traffic data options – incident detection is based on floating vehicle speed data from HERE service. Following sections in this chapter studies vehicle speed data traffic incident detection methods.

3.2.2 Time Series Analysis Methods

In time series based detection methods, the traffic profile for a given road section at a given day is estimated using time series model estimation and forecasting. After the reference traffic profile is estimated, incident detection can be performed in multiple ways. A tracking variable is used as in statistical methods. Changes in the tracking variable are monitored, and if its values cross the defined lower and upper bounds by significant amount of consecutive times, an incident is suspected.

In exponential smoothing method the time series model is estimated by inspecting values in history with exponentially diminishing weights. [21] Tracking variable is a sum of all estimation in history divided by current standard deviation. If the absolute value of the tracking variable is too large for sufficient amount of consecutive observations, incident is detected.

ARIMA models [24] are a family of advanced statistical models designed to estimate and forecast time series with stochastic characteristics. ARIMA is composed of two components. AR, or autoregressive part is used for estimating model with linear regression to time series values in history. MA, or moving average part estimates noise in time series based on history. ARIMA is often enhanced with a seasonal component which is applicable to both AR and MA components. With the seasonal component, time series with a periodic continuity can be estimated with better accuracy. Moreover, time series with a continuous trend can be estimated by adding a differential component to the model. Resulting model is often referred to as SARIMA, or seasonal integrated autoregressive moving average model. Generally all variants of said models are, for convenience, simply referred to as ARIMA models. Traffic incidents can be detected using ARIMA models by estimating and forecasting flow levels on a link for upcoming day and analysing the difference of traffic flow profile between forecast and realised values. [19]

Time Series Analysis Methods require that the data has easily recognisable patterns, such as seasonality and a trend. Ideally, this is true for the traffic speed data but in reality these values have heavy fluctuation because of the constant noise made by imperfect hardware. Additionally, the dynamics of the real traffic patterns are very complex and sensitive – even one low-speed vehicle on a single-lane road can impact on the traffic speeds a long way upstream. Therefore, other methods should also be considered.

3.2.3 Clustering Methods

Clustering [23] is an alternative approach in incident detection. Instead of using tracking variables in the time series, it is divided to multiple groups, or, clusters on the basis of different factors. With a sufficiently small group count, incidents can be detected as separate clusters in the time series. The best way to utilise clustering in incident detection is to study error terms in an estimated time series which can be calculated between traffic speeds on the observed day and the mean values of speed on the equivalent days. If the error term is sufficiently large, the data is clustered to separate groups.

There are several approaches to address the clustering problem. In optimization-type clustering, data points are grouped by optimizing a criterion which is proportional to the integrity of the corresponding clusters. The process is iterative in terms of updating the criterion after each re-clustering for optimal value. One typical criterion to be optimized is a distance from centroid point which is estimated for each cluster in every iteration.

In hierarchical clustering, data is first allocated to N clusters, where N is the number of observations. Then, the closest pair of clusters by predefined criterion are merged to a single cluster. This procedure is iterated until the desired amount of clusters are formed. Hierarchical clustering can also be performed in the other direction, by separating clusters from a single cluster.

One problem with clustering methods is that they normally require the desired amount of clusters to be known in advance. When using these methods for incident detection, there can be cases where incidents are non-existent and as a result, only one cluster should be formed. On the other hand with the anticipation of an incident, the amount of clusters should be set higher than one which results in 'forced clusters' on a clearly incident-free day. Several methods for cluster number estimation have been proposed. One approach, which is intended to be used with centroid methods, compares within-cluster squared distances from corresponding centroid m for current iteration $J_1^2(m)$ and next optimally created cluster division iteration $J_2^2(m)$. If the test value

$$L(m) = \left(1 - \frac{J_2^2}{J_1^2} - \frac{2}{\pi p}\right)\sqrt{\left(\frac{n_m p}{2\left(1 - \frac{8}{\pi^2 p}\right)}\right)} \tag{3.1}$$

exceeds the critical value of an normal distribution, null hypothesis of an homogeneous cluster can be rejected and the cluster count is increased. In equation 3.1 n_m denotes the number of observations in cluster m and p denotes the number of dimensions in the data to be addressed. The problem is that for the normal distribution estimation, every calculation for cluster distances must be performed on the data for one and two clusters which is computationally intensive. Therefore, there is a need for a simpler approach.

3.2.4 Alternative Incident Data Sources

Other sources for incident listings must be considered as a reference to ensure correct detection of ground truth incidents from traffic speed data. There are various alternative incident data sources that can be utilised for measuring ground truth completeness. [11] Not all of items in the list below will be used in this thesis, but should be considered when further developing assessment process in the future.

- *Crowdsourcing.* Many traffic applications for mobile phones offer a possibility to submit one's own information about the state of the traffic, including witnessed incidents.

- *Emergency missions.* A list of emergencies related to traffic incidents can be requested from the Finnish Emergency Response Centre Administration (FERCA). While this emergency data is inaccurate with regard to the reported location, it provides the earliest possible timestamps for incidents in Finland.

- *Towing services.* Some vehicle towing services keep record of the towing tasks carried out in the history. While these listings would not include precise incident start and end times or locations, they could be suitable addition to data that represents all incident cases for reporting completeness studies.

- *Road services.* If a road vehicle has a malfunction beyond so that it cannot be driven, certain voluntary and commercial road services can provide assistance for simple repairs. Some of these services keep record of their repair tasks, which could be useful addition to complete incident listings.

- *Traffic cameras.* Traffic cameras offer a way to detect incidents manually from road sections that are monitored. This can be very challenging considering that there are hundreds of cameras across the Finnish road network. On the other hand, incidents detected from the camera feeds are readily validated.

3.3 Quality Assessment Methods

3.3.1 QKZ

Traffic incident messages generally have two fundamental aspects to be assessed: time and location. In quality criteria, these factors are to be analysed separately by timeliness and location accuracy. However, time and space should not always be measured regardless of each other. In the case of traffic incident messages, proper timing is meaningless if location is erroneous and vice versa. [25]

QKZ and QFCD [20] are approaches that combine time and space components into two measures: coverage and error rate. In QKZ, traffic incident message quality is drawn as a form of heat map where horizontal and vertical axes represent time and space respectively. When colour of an area in the heat map is sufficiently red – or, speed is greatly decreased – congestion caused by incident is suspected. Congested areas can be formed on a time-space graph and compared to corresponding traffic incident messages. As both congestion and traffic incident messages can be represented as areas in the graph, a convenient approach for coverage calculation is to study overlapping of the areas. As seen in the figure 3.1, coverage in QKZ is defined as

$$\mathrm{QKZ}_1 = \frac{D}{E}, \tag{3.2}$$

where D is the area of overlapping regions in incident message and congestion areas and E is the size of the entire congestion area. Coverage is not sufficient to explain the spatio-temporal quality of the incident report by itself. Reports that have excessively large spatio-temporal dimensions compared to corresponding

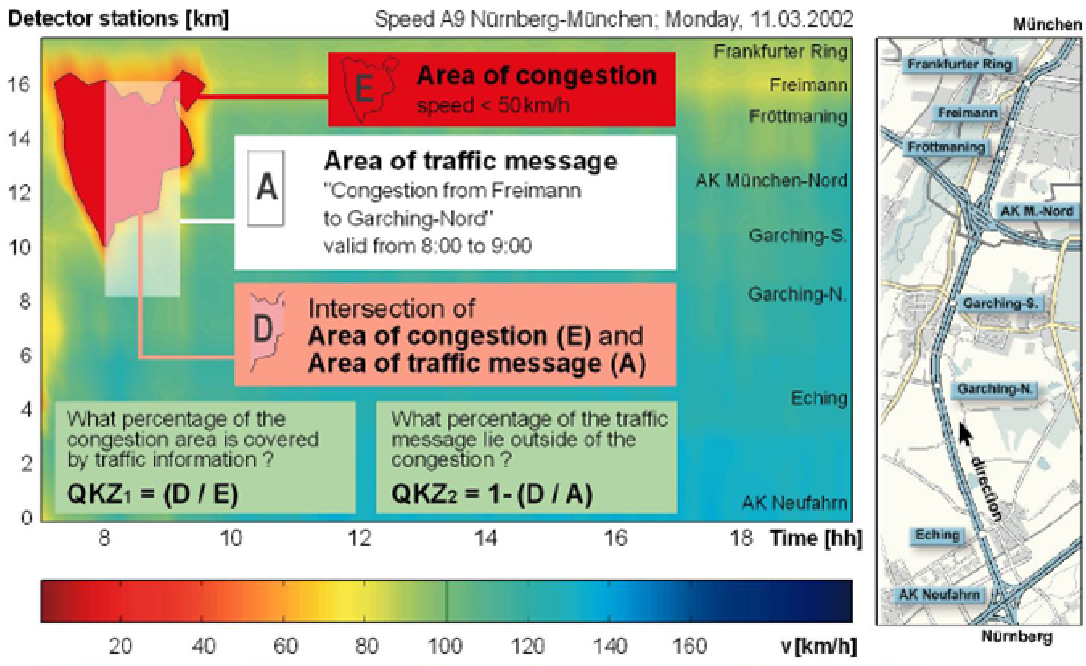


Figure 3.1: QKZ explanation. [20]

incidents are as unwanted as incorrect reports with small dimensions. Therefore, excess message areas are also taken into account as error rate, which is defined as

$$\mathrm{QKZ}_2 = 1 - \frac{D}{A}, \tag{3.3}$$

where A is the area of the incident message. $\mathrm{QKZ}_{1,2}$ quality factors must not be mistaken same as the criteria described by EIP, where coverage and error rate are the proportional amount of records that are correct or erroneous. However it is possible to modify the QKZ method for the aforementioned criteria.

3.3.2 QSRTI

QSRTI [25] is a modification of the QKZ method designed to assess the coverage and error rate in the EIP criteria introduced in section 3.3.1. In QSRTI, safety-related event messages are spanned as an area over a spatio-temporal graph along with ground truth of the corresponding event. Tolerance zones are added to both space and time axes to provide upper error bound for both spatial and temporal criteria. QSRTI areas are illustrated in figure 3.2. Coverage is calculated as

$$\mathrm{QSRTI}_1 = \min\left(\frac{D}{E} + \alpha\frac{\min\left(E, F\right)}{E},\ 1\right), \tag{3.4}$$

where the first term describes the message coverage over actual event and the second term is coverage over tolerance zones. Parameter α represents the relative size of the tolerance zone respect to actual ground truth area. The greater the parameter is, the more the effect of the tolerance zone is taken into account. The

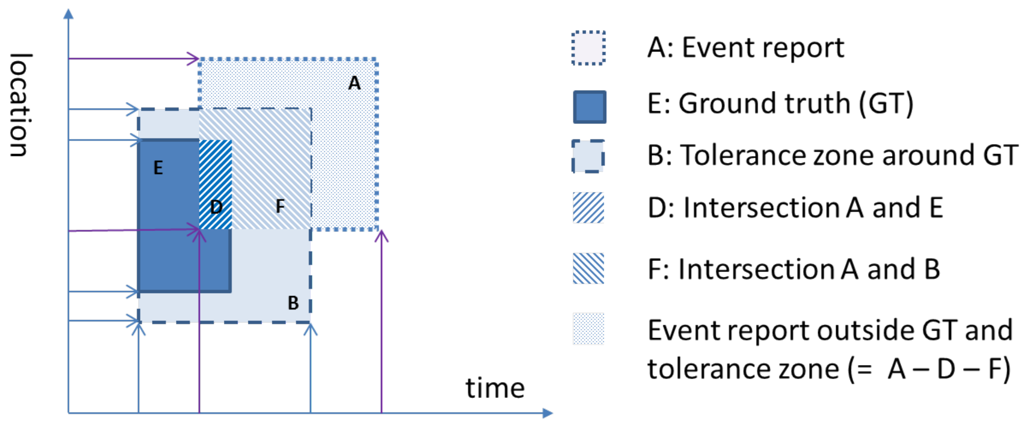


Figure 3.2: QSRTI ground truth and message areas. [25]

minimum function ensures that an excessively large message area is not beneficial for event coverage.

The main differences between QKZ and QSRTI are

- *Categorisation.* Events are divided into categories based on the type of the event. If categories do not match, event coverage is set to zero. Otherwise, area calculations are performed.
- *Tolerances.* Ground truth is expanded with tolerance zones to address problems associated with 'best effort' quality levels in timeliness and latency.
- *Result calculation.* Result is calculated as an arithmetic mean over a sufficient sample of events from all categories.

3.4 Statistical Process Control

Statistical process control is a set of methods in which a quality of the process is measured using statistical tools. Commonly utilised seven quality tools introduced by Karou Ishikawa [29] are used to measure, monitor and control manufacturing processes. These tools are originally designed to be used in manufacturing industry but are also suitable for measuring the quality of the incident reporting because of their general nature.

Potential causes for process quality variations are outlined in a cause and effect diagram (also known as Ishikawa Diagram or fishbone diagram). Commonly, four main causes are listed in the diagram: man, machine, material, and method but variations can be performed depending on the context. These factors are usually further divided into sub-factors for more granulated overview of the components affecting process quality.

Process quality problems are documented on a check sheet with hash marks on a predefined list of typical defect cases along with a description about the problem. These records can be used on a statistical analysis, such as studying defect distributions over different types of events.

Relationship between different process variables can be visualised with scatter diagrams. For example, mutual reactivity for quality between variables can be analysed by calculating correlation or covariance for the corresponding datasets or by finding same patterns graphically in the diagram.

Flow chart representation is an important part of the quality assurance to understand the underlying complexity of road incident reporting and quality-critical parts of the process. The chart is typically drawn as a sequence of required steps in a process to achieve the desired result.

Overall frequency for different types of quality problem causes can be visualised with a Pareto diagram. Basically, Pareto diagram is a histogram of various quality problem-causing events that sorted in decreasing order of the event frequencies usually accompanied by a cumulative probability curve.

Drawing histogram over different frequencies of problem causes or other interesting events in the incident reporting process is a good practice and helps in process outlining.

With a control chart, an operator can monitor if the incident reporting process is in control, i.e. all quality variables stay in acceptable limits at all times. Control charts can be drawn from individual quality measures or more general, composite indicator can be formed as a function which depends on the individual quality values. Process stability is monitored by defining upper and lower control values (UCL and LCL, respectively) and ensuring that the quality indicator stays within defined control limits. The process is out of control if sufficient amount of values fall outside of the control limits or indicator values are systematically larger or smaller than mean indicator value.

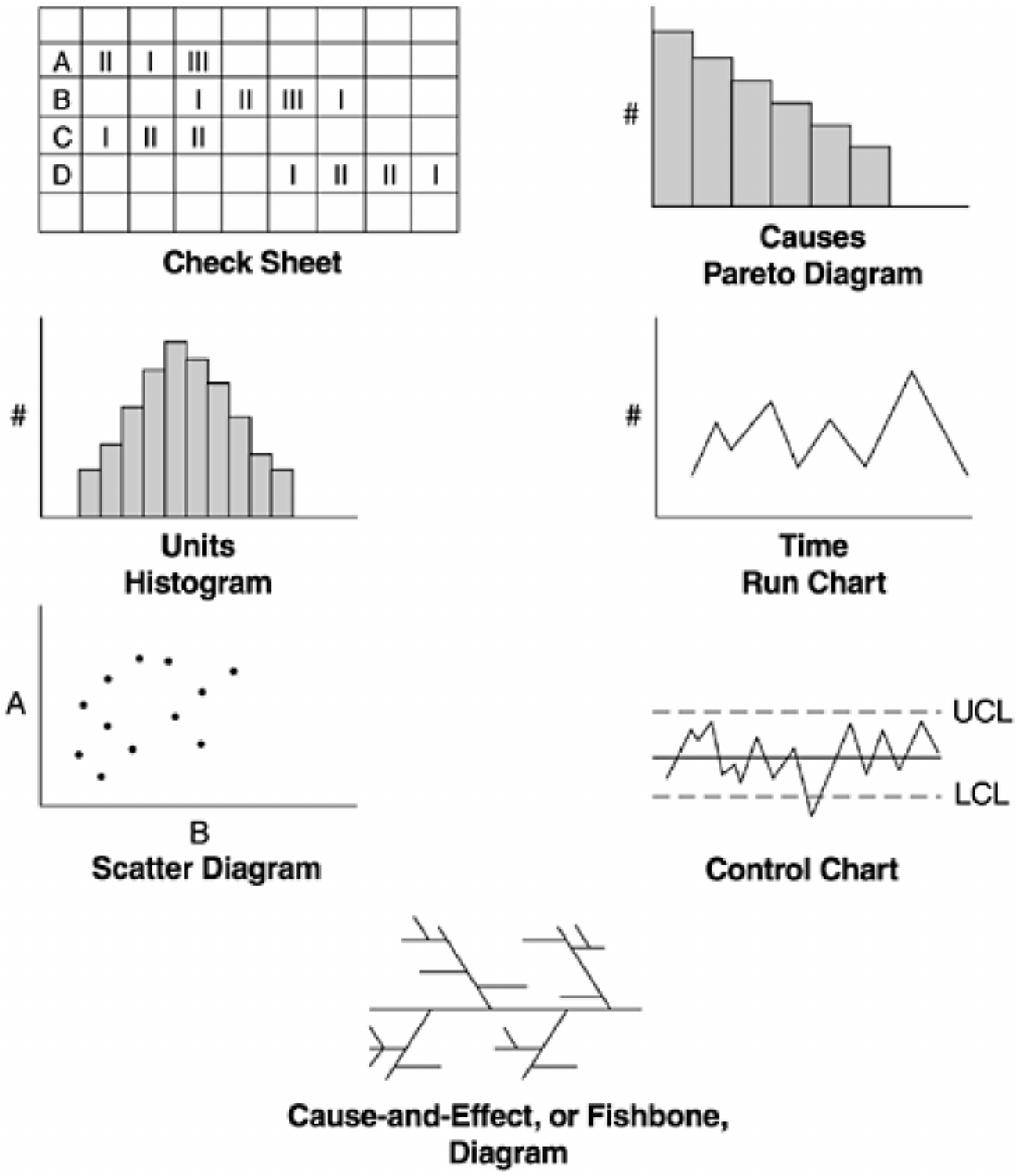


Figure 3.3: Ishikawa seven basic tools of quality. [28]

Chapter 4

Methods

4.1 Speed Distribution Tests

Following the problems associated with the incident detection methods in chapter 3, a new method is proposed. In this method, every HERE measurement link in every day at 6 AM–9 PM in a study period are processed by comparing the traffic profile on the studied incident day to all reference days on compatible weekdays. For every studied day on a link, excessively peaked speed values in a time series are first filtered using a two-sided moving average from two values on both sides. Reference days are chosen from the following categories based on the assumption for similar traffic profiles on each group:

- Monday–Thursday,
- Friday
- Saturday
- Sunday

Reference speed values are required to adhere a certain distribution for acceptable detection performance. After the collection of reference day samples the corresponding values for every studied day time series values in a 15-minute window are filtered. These values are assumed to be normally distributed. In figures 4.1, 4.2, and 4.3 reference speeds in Helsinki, Kotka, and Jyväskylä from morning rush hours, middays and afternoon rush hours on a randomly picked weekday are drawn in histograms along with the estimated normal distribution density curve. In Kotka and Jyväskylä the reference speeds are seemingly normally distributed. On the other hand in Helsinki the speed distributions seem to be much more complex. At the morning rush hour and at midday the distributions are skewed to the left but are otherwise relatively normally distributed. At the afternoon rush hour the distribution is slightly skewed to the left and is slightly deformed for a normal distribution. The rush hours in Helsinki capital region Ring I are focused at the afternoons when eastward traffic is studied as typical direction of the work commute is to the west. Therefore, the normality of the reference speeds suffers from the excessive amount of traffic and speed declines.

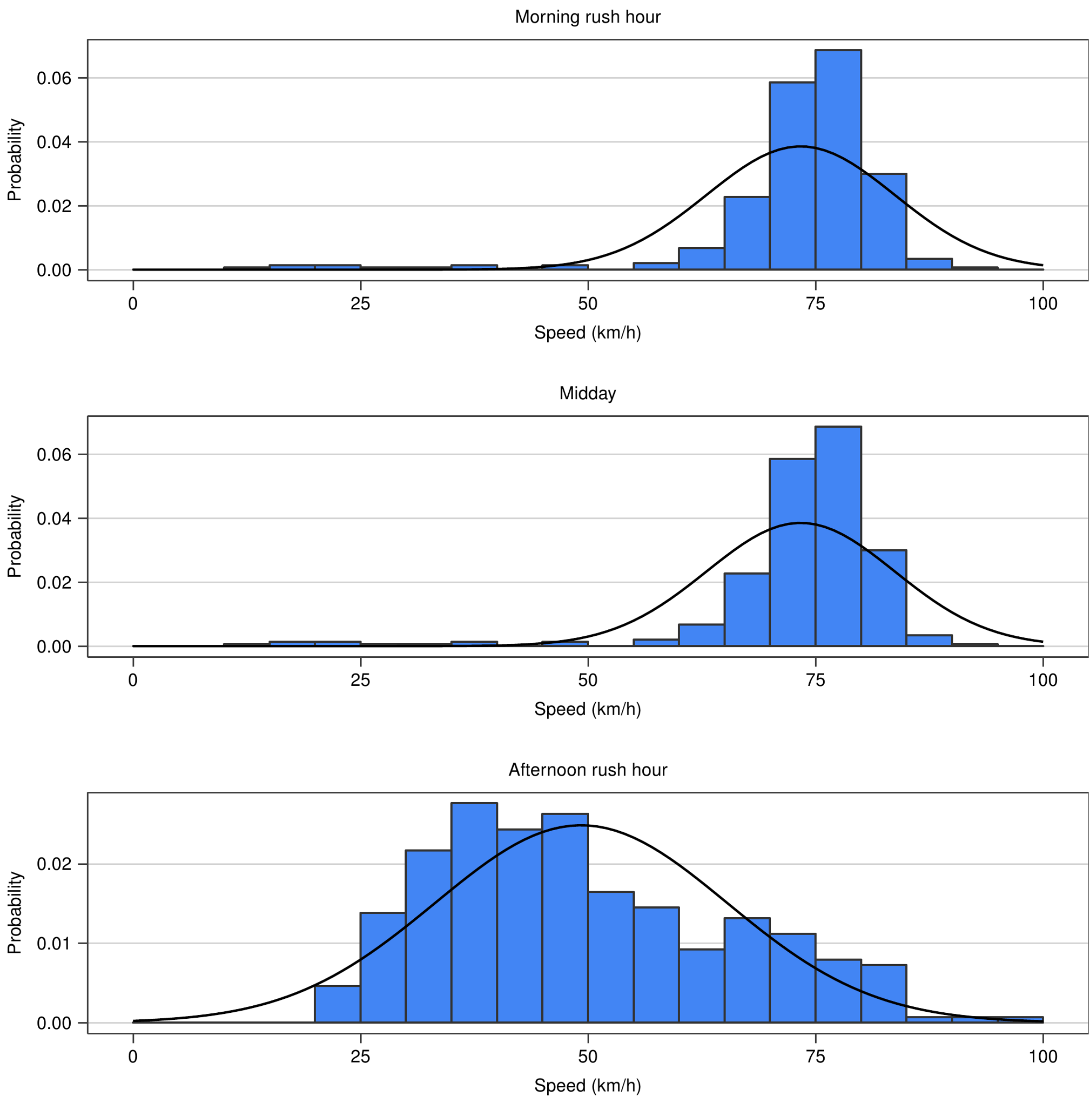


Figure 4.1: Histogram and density curve from reference speeds at Ring I section in Konala, Helsinki. Direction is to the east.

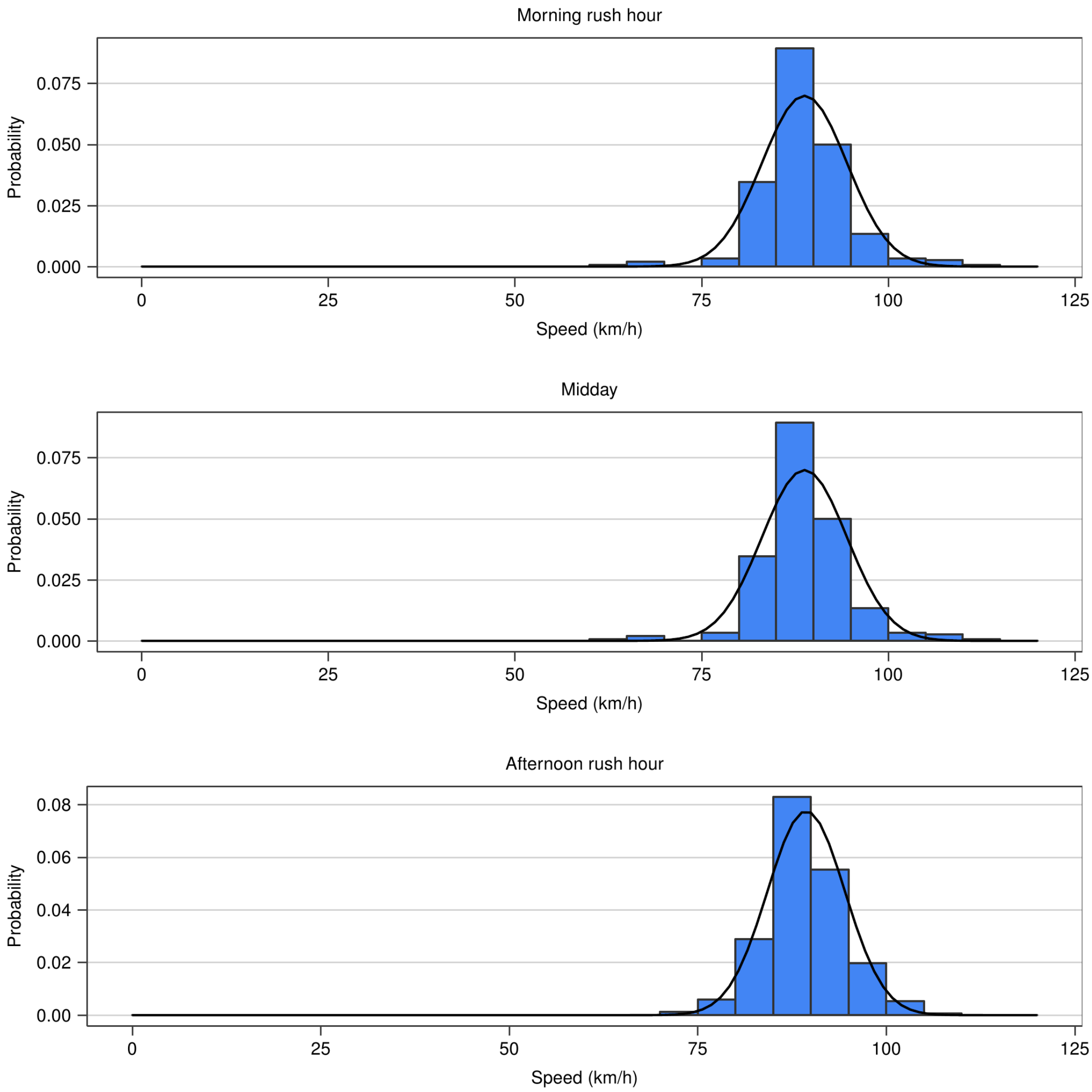


Figure 4.2: Histogram and density curve from reference speeds at Highway 4 section in Jyväskylä. Direction is to the north.

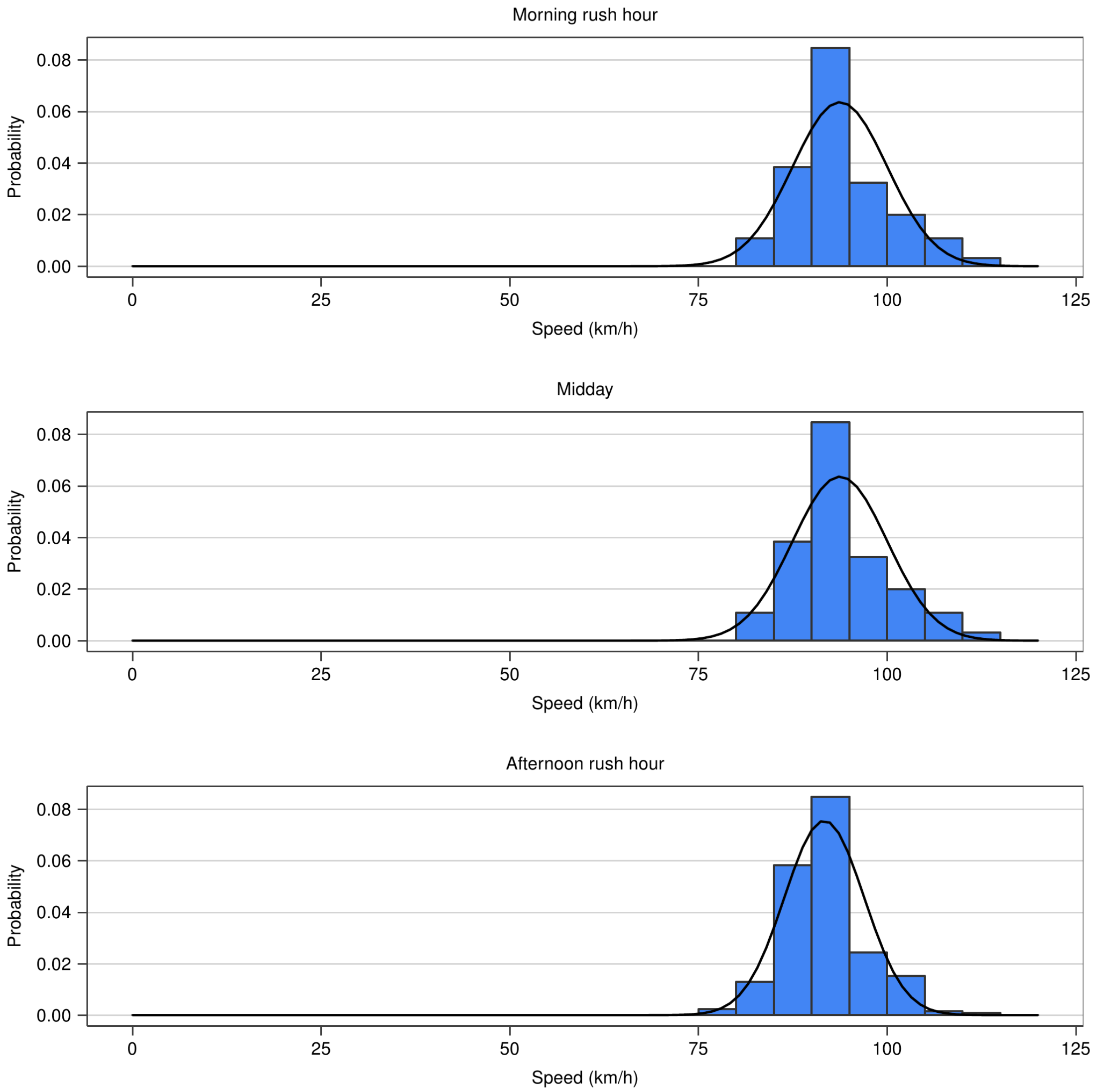


Figure 4.3: Histogram and density curve from reference speeds at Highway 7 section in Kotka. Direction is to the east.

The normality of the reference day speeds is studied by performing two statistical tests for a randomly-selected set of time windows from the whole dataset. Both tests, Kolmogorov-Smirnov and Anderson-Darling, are based on the comparison of the reference cumulative distribution and an empirical distribution function (EDF) which is defined as

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_x(X_i), \tag{4.1}$$

where X_i are independent identically distributed, randomly sampled values from the data, and $I_x(X_i)$ is an indicator function, which is 1 if $X_i < x$ and 0 otherwise. In the Kolmogorov-Smirnov test, the maximum distance between the empirical cumulative distribution and the reference cumulative distribution is calculated. [30] For ordered data points $x_1 < x_2 < x_3 < ... < x_n$ the test value is

$$T = \sup_x |F^*(x) - F_n(x)|, \tag{4.2}$$

where $F^*(x)$ is the reference cumulative distribution function. If the test value T is greater than the critical value from the Kolmogorov test statistic on a significance level $1 - \alpha = 0.95$, the null hypothesis for the currently studied distribution adherence is rejected. The Anderson-Darling test statistic for the same ordered data points as in the Kolmogorov-Smirnov test is defined as

$$W_n^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left(\log F^*(X_i) + \log 1 - F^*(X_{n+1-i})\right). \tag{4.3}$$

If the test value W_n^2 is greater than the critical value from the Anderson-Darling test statistic, the null hypothesis for distribution adherence can be rejected.

The distribution of choice is robustly estimated on given values using maximum likelihood estimation, and 95 % and 60 % confidence intervals are calculated for each reference value group. With robust estimation, time series on the reference days need not to be incident-free.

Possible normal distribution compliance for reference speed values was tested with mentioned tests. A $n_{inc} = 50$ sample was collected from all delimited Digitraffic incidents. In every sampled incident $n_t = 50$ points of time were sampled. This resulted in a $n = 1898$ sample of reference value sets, because some of the incidents had an incomplete time series with insufficient amount of time points for a desired sample size. A distribution compliance is presented in the table 4.1. The normal distribution has relatively large share of acceptance for null hypothesis H_0 for distribution compliance. Based on the results, normal distribution is acceptable for the analysis.

4.2 Incident Detection

Incident detection sensitivity is calculated from a sample of reports from Digitraffic. For every incident, the time series from the studied incident day is loaded along with the estimated 95 % and 60 % confidence intervals from the reference

Table 4.1: Distribution test results.

Test Name	H_0 Acceptance
Kolmogorov-Smirnov	88.6 %
Anderson-Darling	74.3 %

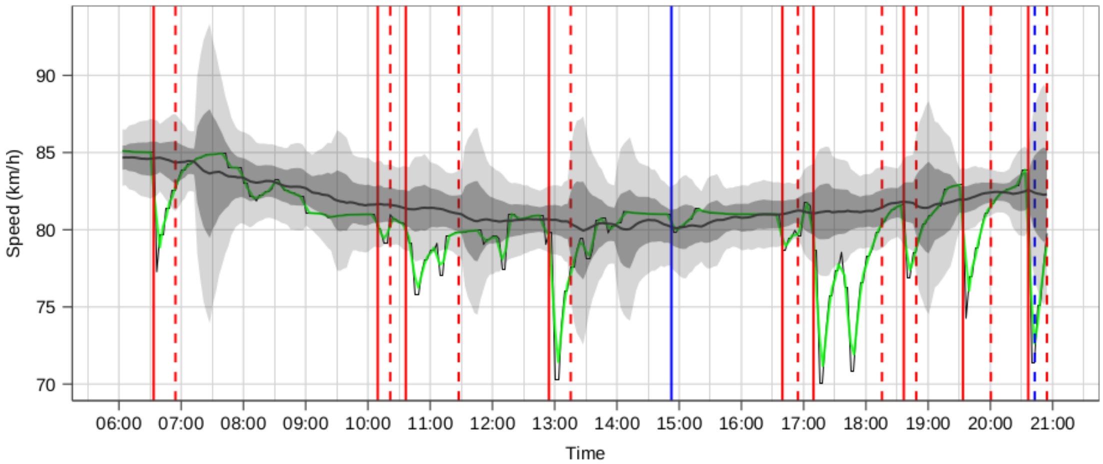
days. The relative amount of values that are lower than the 95 % confidence interval in a traffic incident message are calculated as well as the time of the day when the incident occured. Mean shares of speed values violating the confidence intervals are calculated for every hour of the day. The consistent normality of speed values in reference days discussed in section 4.1 is required for robust estimation of detection sensitivity parameters.

The time series speed values on each day in the study period and all HERE measurement links in the delimited network are analysed with a chronologically moving 10-value time window. If at least the previously estimated relative amount of speed values in a time window have smaller values than the lower 95 % confidence interval edge, an incident is suspected. The start and end times for incidents are defined as the first and last points where the studied time series speed values are below the 60 % confidence interval.

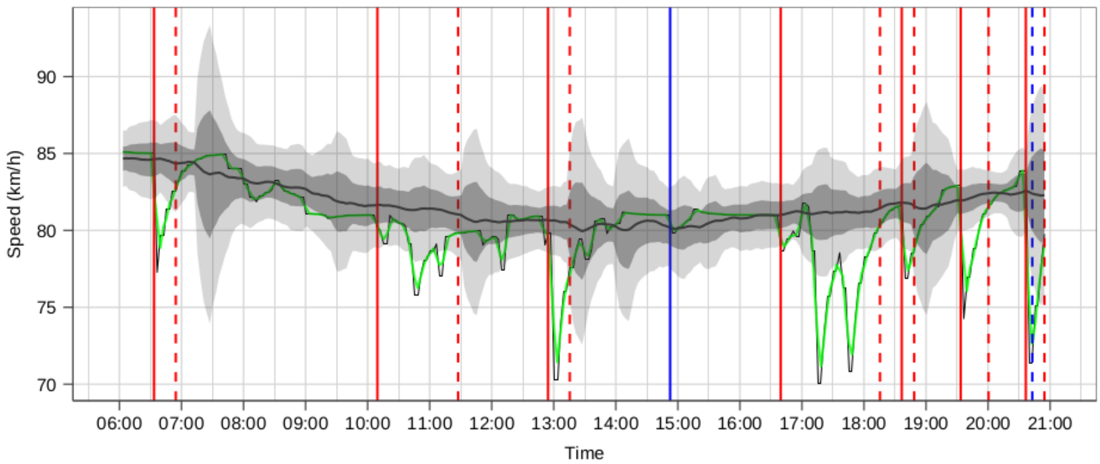
Following the excessive peaks in the speed values in the time series – even after the moving average smoothing – an excessive number of incidents are detected as clusters in the vicinity of a suspected incident. These incident clusters are combined to one incident if the neighbouring end and start times in a cluster differ less than 15 minutes. For sensitivity analysis, non-combined incidents are also saved. An example of the non-combined detected incident clusters is in figure 4.4a. In figure 4.4b four incidents are combined into two incidents thus decreasing the amount of incidents clusters in the time series.

Traffic incidents have complex, dynamic impacts on traffic flow and speeds on both upstream and downstream road sections. Therefore, incidents have to be not only combined temporally but also aggregated spatially. This is performed by studying links adjacent to the current incident links and combining adjacent incidents iteratively if

- incident end time on an adjacent upstream link is later than the respective incident end time and start time is between the start and end times on a current link or
- incident start time on an adjacent downstream link is earlier than the respective incident end time and end time is between the start and end times on a current link.



(a) Non-combined incidents.



(b) Combined incidents.

Figure 4.4: Speed time series with non-combined and combined detected incidents. Solid and dotted red lines represent respective start and end times of the detected incidents and blue lines represent reported incident start and end times.

4.3 Combining Data

After data collection and incident detection the dataset in the study includes incident records from the FTA, detected incidents from HERE speed data, and emergency records from the Finnish Emergency Response Centre. Two data sources are used for every criterion.

- *Timeliness:* HERE and FERCA
- *Latency:* Digitraffic and FERCA
- *Location accuracy, error rate and event coverage:* HERE and Digitraffic

Data is joined spatio-temporally and maximum tolerances are needed to filter the possible outliers. These tolerances are set manually by testing different values. Data from every source is first fixed to HERE network geometry by calculating distances from the incident coordinates to the HERE measurement links which consist of several shorter and linear sublinks. Distance is calculated as a minimum value over all sublinks. Measurement link with the minimum sublink distance to the incident is selected.

In temporal data joining the time tolerances are set based on the sources. When joining incidents detected from HERE and Finnish Emergency Response Center data for the timeliness criterion, all incidents from the HERE in the range of $[t_h - 30, t_h + 30]$ minutes are studied where t_h is the detected start time of the incident. The correct HERE incident is selected by finding the minimum absolute time difference to the FERCA incidents in the mentioned range.

For the latency criterion all FERCA emergency missions are studied in the range of $[t_d - 60, t_d]$ where t_d is the time of the preliminary incident announcement in the Digitraffic data. The range is based on the assumption that a Digitraffic record is not created before the corresponding FERCA record.

For location accuracy, error rate, and coverage criteria Digitraffic data is joined with the detected incidents from HERE speed data. In this operation, all detected incidents in the Digitraffic timestamp range $[t_d - 30, t_d + 30]$ are studied and the incident detected with the minimum absolute time difference is selected.

4.4 Method Testing

From the criteria in table 1.1, timeliness (start), location accuracy, error rate and event coverage are assessed. Both error rate and event coverage criteria values are calculated with two variants: first by the EIP definition whether the incident report covers the actual incident at all and second with a spatio-temporal area calculation method introduced in section 4.6. Although the area calculation method for error rate and event coverage criteria also implicitly assesses timeliness and location accuracy criteria combined, they are also studied separately because of their respective quality level specifications in EIP criteria list.

Timestamps for emergency missions are the closest available time values in relation to the incident detection times in FTA incident reports, because the time

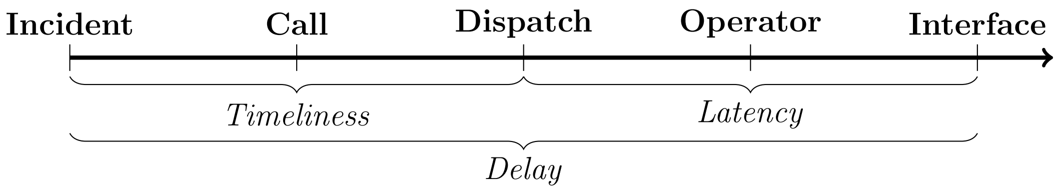


Figure 4.5: Timeline of the incident data formation and temporal criteria sources.

of arrival for incidents in the incident queue of the HäTi system are not logged. Therefore, these timestamps are used for the calculation of the timeliness (start) criterion. Timeliness is calculated by subtracting timestamps in the Emergency Response Center timestamps with the corresponding start times measured from the ground truth. In some cases, end times of the traffic incidents are not declared. This will be addressed by defaulting the end time stamp to the specified time-out value of 90 minutes in the FTA incident information specification after the last update of the event.

The latency criterion can be seen as a systematic error between the time when the incident is detected and the moment when the information reaches the DATEX II interface in the Digitraffic systems. The latency criterion value is calculated as a time difference between these two phases from the FERCA and Digitraffic timestamps. The timeline of the preliminary incident detection and reporting process with the temporal criteria calculation points are depicted in figure 4.5.

Quality level of location accuracy in EIP criteria is defined not only by location error, but also as a correctness of link between intersections in lower quality levels. Therefore, location accuracy is calculated as euclidean distance between the ground truth and the message, and as a truth value whether the location is reported on the link where the incident originally occured. For point-type locations, traffic incident locations from FTA are set to a corresponding HERE measurement link. This way, spatial accuracy is calculated as the mean value from both ends of the reported road section in relation to the equivalent ground truth measurement links detected from HERE.

Coverage and error rate criteria are calculated using a method modified from for real-time incident information. In the measurement, both time and location values are the same as the ones used in the timeliness and location accuracy calculation. Both ground truth and incident reports to be assessed are placed on a time-location space. Then, coverage and error rate criteria are calculated as a ratio of overlapping and outlying areas. In QSRTI, ground truth is augmented with tolerance areas in both time and space dimensions. The reason for this is that QSRTI is designed to address location accuracy, timeliness, coverage and error rate criteria altogether and lower quality level requirements for 'link between intersections' in location accuracy and 'best effort' in timeliness are presented as the mentioned tolerance zones. These tolerances, however, are not used in method testing in this thesis because location accuracy and timeliness are also assessed separately in addition to the modified QKZ testing for coverage and error rate. A modified method based on QKZ is introduced in section 4.6. Additionally, the event coverage criterion is calculated as a strict value whether the incident message is located in the most downstream HERE measurement link at the right moment in time.

Criteria scores are calculated for every incident case. With these criteria scores, overall level of incident reporting quality for Finland is defined according to the table 1.1. Additionally, incident reports without the timeout value of 90 minutes are studied separately for both variants of error rate and event coverage criteria and associated quality levels are calculated. Finally, the latency criterion is analysed for incident reports with and without preliminary announcements.

4.5 Defining the Level of Quality

Results from the measurement-based criteria definitions and custom methods are interpreted, and overall level of traffic incident reporting quality in Finland is estimated based on these results. The definitions for different levels of quality in every criteria is shown in table 1.1. For timeliness, latency, and location accuracy the level of quality is chosen whether at least 95 % of the calculated values are lower than the written value levels. For error rate and event coverage the percentage is directly compared to the defined value levels. Based on the results, recommendations are proposed for incident reporting quality optimisation. Also, possible ideas for criteria refinement are documented over the course of the method testing.

4.6 QRTTI

Both event coverage and error rate quality criteria are calculated as a summarised fuzzy value over all incidents in a specified time frame. These values, named QRTTI_1 and QRTTI_2 respectively, are measured as a relative ground truth and incident message area sizes.

An imaginary traffic incident case is drawn in a spatio-temporal graph in figure 4.6. In the figure, A (blue) is a detected incident ground truth area and B (red) is a FTA traffic incident message area. Both the ground truth and the incident message evolve over time as the speed decreases and increases in the upstream traffic links, and the FTA incident reports react to the situation with a delay. QRTTI_1 is defined as fraction of ground truth area where the incident message is valid

$$\mathrm{QRTTI}_1 = \frac{f(A \cap B)}{f(A)}, \tag{4.4}$$

where $f(.)$ is an area calculation function. QRTTI_2 is defined as a fraction of the message area not covering the ground truth

$$\mathrm{QRTTI}_2 = \frac{f(B) - f(A \cap B)}{f(B)}. \tag{4.5}$$

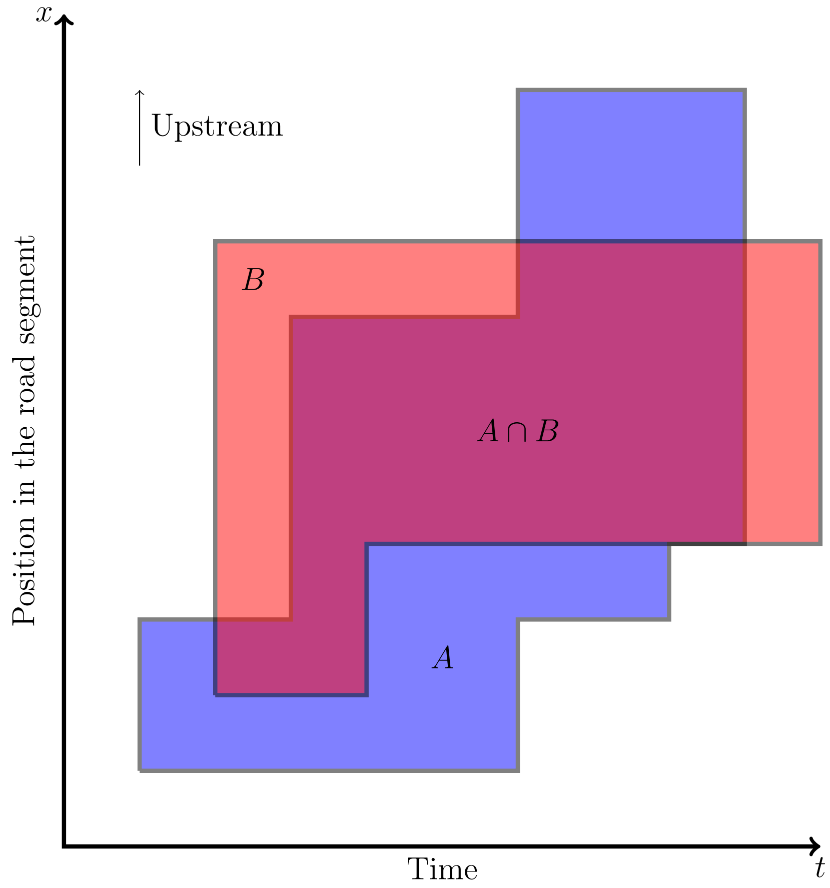


Figure 4.6: QRTTI.

The coverage C and the error rate E of the entire study area in a study period is calculated as the arithmetic mean over all corresponding measures with all N number of traffic incidents

$$C = \frac{\sum_{i=0}^{N} \mathrm{QRTTI}_1}{N} \tag{4.6}$$

$$E = \frac{\sum_{i=0}^{N} \mathrm{QRTTI}_2}{N}. \tag{4.7}$$

4.7 Quality Assurance

The EIP(+) project obliges the participating countries to integrate quality assurance process into operative functions. Fundamentally, this means that the compliance of the defined quality levels has to be monitored in real time.

Quality assurance can be performed by using a control chart as stated in section 3.4. Compound value calculated from the event coverage and error rate criteria during the past week is used as an indicator

$$I = \frac{1}{7}\sum_{i=1}^{7} \frac{\mathrm{QRTTI}_{1i} + (1 - \mathrm{QRTTI}_{2i})}{2}, \tag{4.8}$$

where QRTTI_{1i} and QRTTI_{2i} are average event coverage and error rate values respectively from day i. The indicator value is updated daily. In the equation, both criteria have equal weights while additional parameters can be added to distribute the attribute weights.

With control chart, indicator values are monitored and the operator is alerted if the value is outside predefined thresholds. These thresholds are referred to as control values. The lower control value (LCL) and the upper control value (UCL) are defined according to the assurance process objectives. On certain cases, both control values are not necessarily needed. One example method is to set values according the EIP criteria level of quality value definitions. For example, if the traffic incident reporting is at special four-star level in both error rate and event coverage, only LCL is adjusted. The control value is defined as

$$\mathrm{LCL} = \frac{\mathrm{LCL}_{EC} + (1 - \mathrm{LCL}_{ER})}{2} = \frac{0.9 + 0.98}{2} = 0.94, \tag{4.9}$$

where LCL_{EC} and LCL_{ER} are the lower value limits of event coverage and error rate in the four-star quality category. If the indicator value I drops below the LCL, two choices are possible: either

i) execute corrective actions to the incident reporting or

ii) if the LCL violation is permanent and short-term corrective action is impossible without major structural changes in the reporting process, lower the quality level rating.

In addition to the lower and upper control values, a central control limit (CCL) can also be defined. This value is usually a long-delay moving average over the indicator values in the history. With the central value, minor systematic errors in the process can be detected when the indicator value is constantly greater or lower than the central value.

In figure 4.7 is a four-field presentation of the event coverage and error rate values. The colour gradient represents the indicator value – the lowest value is in red and the highest value is in green. In field A are the criteria values that have high event coverage and error rate values and thus average indicator scores. This means that the incident report covers excessively large area around actual incident

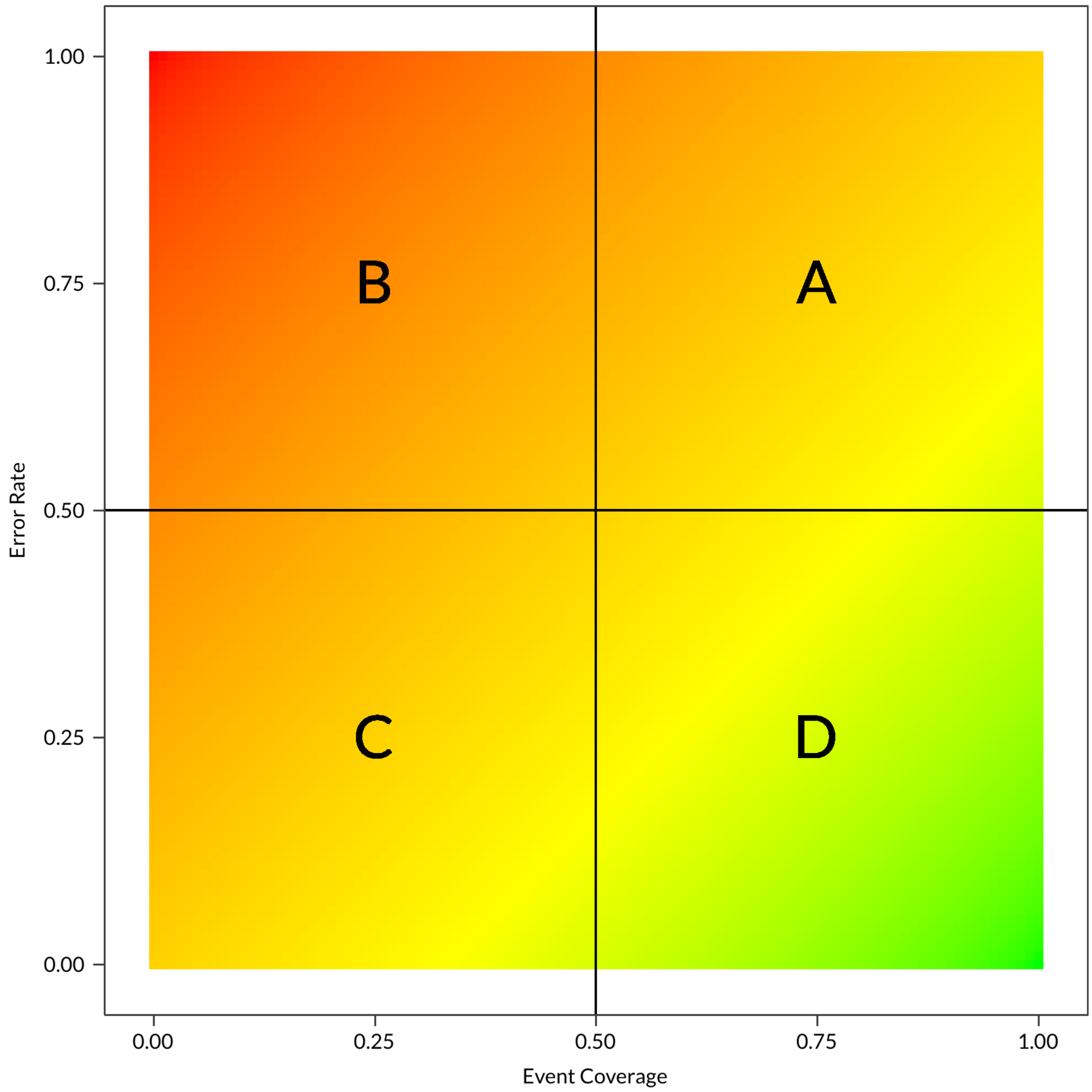


Figure 4.7: Four-field graph of the event coverage, error rate, and indicator values.

area. In field B are the values which have low event coverage values and high error rate values causing the indicator to have minimal values. Thus, the field represents the incident reports that largely or completely miss the actual incident area.Field C has average indicator values as both the event coverage and the error rate are low which means that the field represents incident reports that mainly cover the actual incident area without being too large. The highest indicator values are calculated from the criteria values on the D field where the event coverage values are high and error rate values are low. These incident reports cover the incident area with minimal overlap outside the area.

The four-field presentation can also be used for the quality assurance by placing individual incidents on the graph based on their calculated error rate and event coverage values. This provides a convenient way to check if the criteria values tend to stack to a specific corner or edge on the graph and thus have tendency for specific combination of the criteria values.

4.8 Structural Quality

The proper use of reporting forms is important for the optimal utilisation of real-time traffic information. In addition to assessing the quality of incident information, the structural quality and adherence to the DATEX II schema is studied by manually inspecting several reports and finding possible problems.

Chapter 5

Results

5.1 Incident Detection

The number of incidents detected was over $95,000$. This means that the incident detection algorithm is very sensitive up to the point where various declines in the HERE speed data values not tied to any incidents are consistently depicted as incidents. While the number of detected incidents is high, the distribution of the incidents over all roads were approximately the same as in the Digitraffic reports. This can be seen in figure 5.1 where the percentage of all detected incidents and Digitraffic incident reports per kilometre over all studied roads are compared against each other. Therefore, it is safe to assume that the incident detection algorithm is not detecting incidents at random.

5.2 Data Combination

Initially, Digitraffic records consist of 388 unique incident cases. When the Digitraffic data is combined with the HERE data for timeliness calculation, the filtered number of incidents is 213. After combining the Digitraffic data with the Finnish Emergency Response Centre data for latency criterion calculation, the number of incidents was filtered down to 192. When calculating the location accuracy, error rate, and event coverage criteria, the data is joined with the incidents detected from HERE and the resulting number of cases are 125 for location accuracy and 124 for error rate and event coverage. The overlap of the incident data sources are depicted in figure 5.2. While all datasets lose between 38–68 % of the incident information, the sample sizes are adequate for the calculation of the criteria values.

While the size of ground truth samples are sufficient for the quality assessment, it is not certain that they accurately represent the studied incidents. In figure 5.1 percentage of incident density per kilometre over all roads is presented. The figure suggests that the detected incidents are identically distributed on different highways and roads when comparing to the Digitraffic incidents.

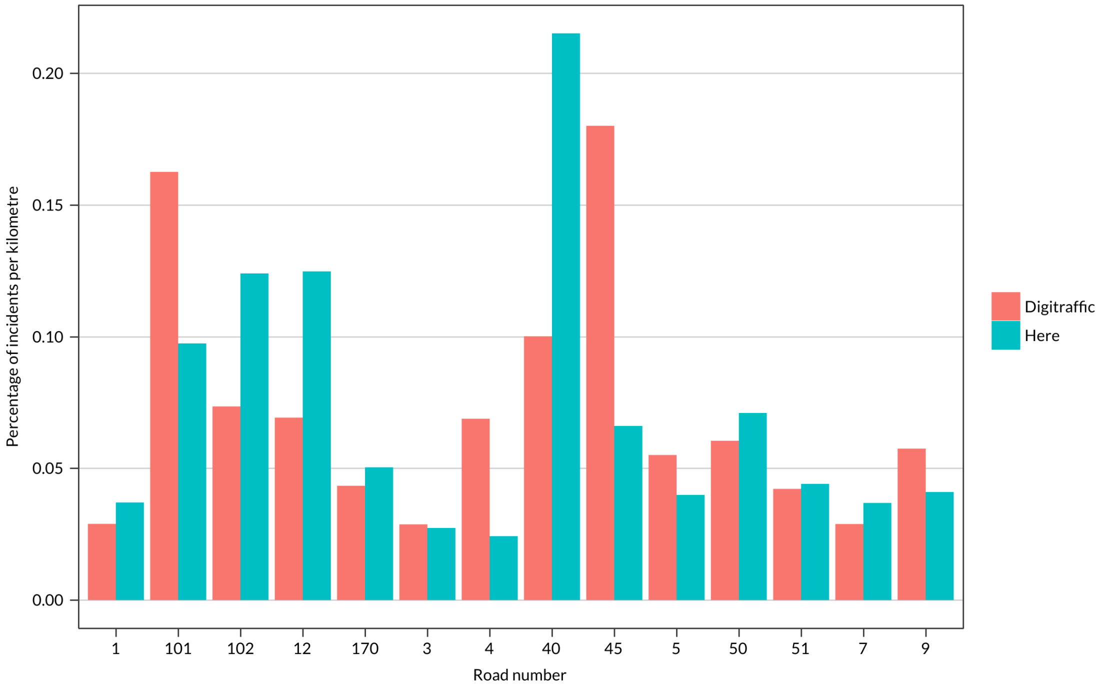


Figure 5.1: Percentage of incident numbers per kilometre on all studied roads.

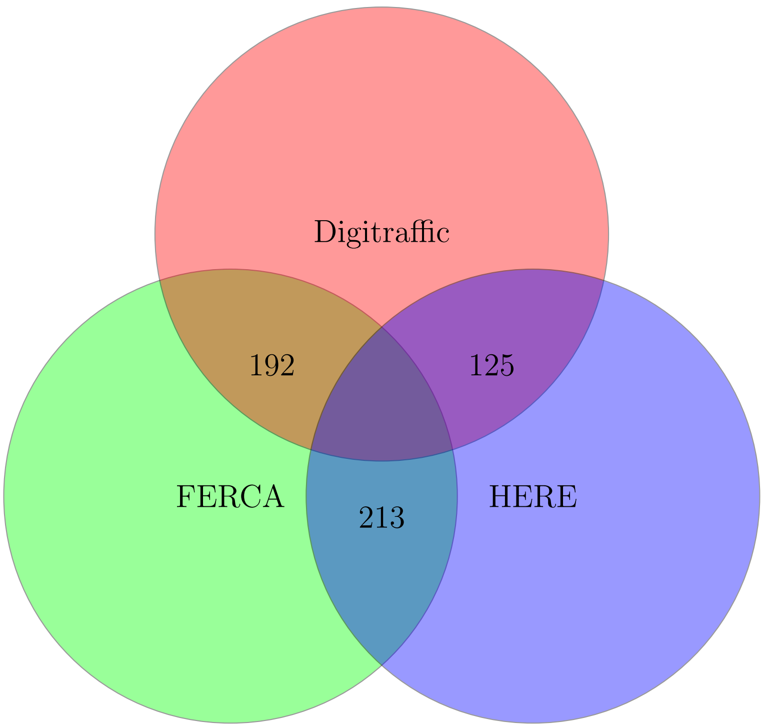


Figure 5.2: Overlap of the incident data sources.

5.3 Quality Assessment

All criteria values were first calculated as a mean and median values to analyse the quality of the incident information numerically. Additionally, traffic incident reports without a default timeout of 90 minutes were analysed separately to study the possible effects of the timeout usage. Median values are presented in table 5.1.

In table 5.1 final results for the criteria are presented along with the proposed level of quality interpreted from the table 1.1 with principles introduced in section 4.5. In table 5.2 are values from both EIP definition and QRTTI variants of the error rate and event coverage for non-timeout incidents. The distributions of the criteria values are presented in histograms in appendix C along with density curves estimated with the maximum likelihood method.

Approximately 80 % of the incident reports have preliminary announcements. The average latency criterion value for the mentioned incident cases is 4 minutes 15 seconds and the median value is 3 minutes 4 seconds. These values are somewhat lower than respective values for all incident cases. Again, the 95 % quantile is 10 minutes 34 seconds which is significantly lower compared to all cases. This means that outlying values in the 20 % share of cases with no preliminary announcements have great impact on the latency criterion.

The share of incident cases where the latency criterion value is over 10 minutes is 20 %. Over 75 % of these cases have no preliminary announcements.

5.4 Quality Assurance

Because the measured event coverage and error rate criteria values are too weak for assigning control values along with the principles introduced in the section 4.7, control card testing is performed with manually defined control values that provide possible violations.

In figure 5.3, the indicator values per week in the whole study period are drawn with a blue line. Control values LCL = 0.45 and UCL = 0.75 are drawn with a red line as the limits of the accepted values for the indicator values along with the central control value CCL = 0.54 as a green line. As can be seen in the figure, most of the indicator values reside between the control limits, with the exception of few cases. However, these cases will not require any immediate action as their successive values return to the controlled limits within a few steps.

Here, the central control value is calculated as the arithmetic mean over all indicator values in the data. The figure shows that the indicator value varies on both sides of the central control value. This means that there are no systematic errors in the incident reporting and the process is in control.

In figure 5.4 indicator values for individual incidents are placed on the four-field graph introduced in chapter 4. In the figure incidents appear to be spread to all fields relatively evenly. There are three exceptions. First, 14 incidents have the weakest indicator value. This means that the incident does not cover the ground truth incident area at all. Second, several values are stacked to the upper-right edge of the graph. This indicates that a large share of the incidents cover the ground truth completely but have a great overlap in the area and thus have too

Table 5.1: Median criteria values and the level of quality according to table 1.1.

Criterion	**Sample Size**	**95 % Quantile**	**Mean**	**Median**	**Level of Quality**
Timeliness	213	27 min 55 s	13 min 45 s	13 min 29 s	⋆
Latency	192	24 min 38 s	6 min 55 s	3 min 38 s	–
Location Accuracy	125	6 km	1.6 km	0.5 km	⋆⋆
Error Rate (EIP def.)	124	100 %	13 %	0 %	⋆
Event Coverage (EIP def.)	124	0 %	87 %	100 %	⋆
Error Rate (QRTTI)	124	100 %	57 %	66 %	–
Event Coverage (QRTTI)	124	0 %	62 %	66 %	⋆

Table 5.2: Applicable median criteria values and the level of quality for non-timeout incidents according to table 1.1.

Criterion	Sample Size	95 % Quantile	Mean	Median	Level of Quality
Error Rate (EIP def.)	55	100 %	7 %	0 %	⋆⋆⋆
Event Coverage (EIP def.)	55	0 %	93 %	100 %	⋆⋆⋆⋆
Error Rate (QRTTI)	55	98 %	42 %	37 %	–
Event Coverage (QRTTI)	55	4 %	57 %	59 %	⋆

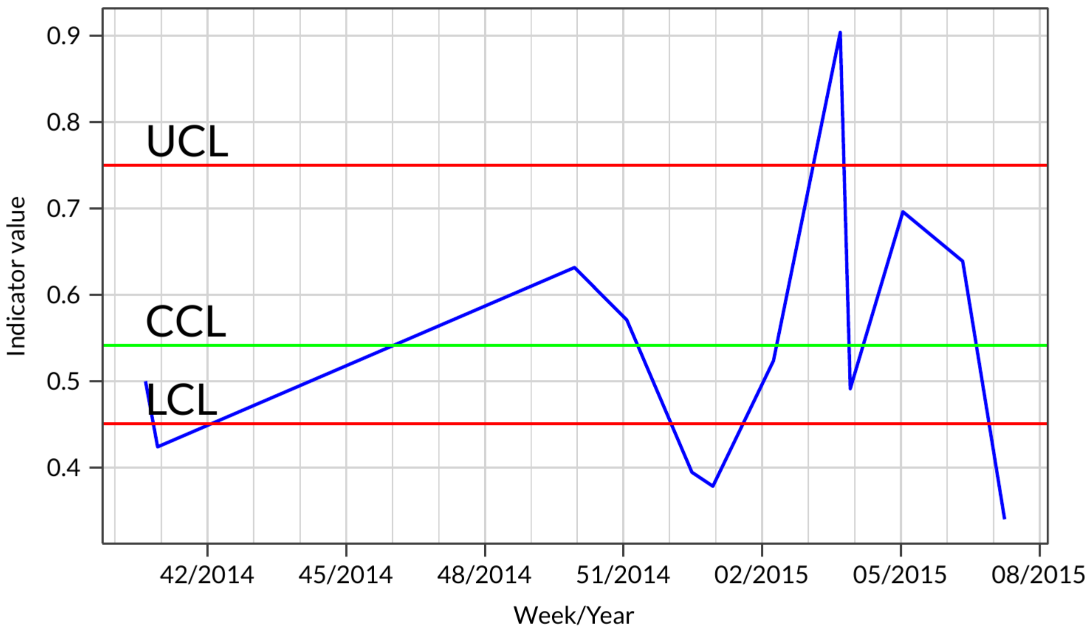


Figure 5.3: Control card for the traffic incident quality assurance.

large incident report areas. Last, several incidents are placed on the bottom of the graph and scattered evenly by the vertical axis. This suggests that there are multiple incident cases where the message area is very small and does not cover the whole ground truth area.

5.5 DATEX II and AlertC Quality

The incident information provided in the Infotripla interface is in DATEX2 format. The location and description of the incident is referenced with AlertC standard. Proper parsing of the information requires that the DATEX2 format and AlertC location reference guidelines are followed properly.

The direction of the traffic in which the incident has an effect on the road is very often indicated as “both”, yet few accidents usually affect both directions in traffic. It is evident that the “both” value is the default value for the incident direction.

A large share of the incident types listed in the traffic information are marked with “other” value despite that the description clearly defines the incident type. This was common in typical road crashes. On the other hand in more unique cases – such as a crashed lorry – the incident type is explicitly defined.

The description of the incident in the AlertC section of the incident message has several problems. The readability and automatic parseability of the text is poor because of inconsistent use of spacing. If the description is too long, it is split into two different values that have to be detected from the structure and concatenated.

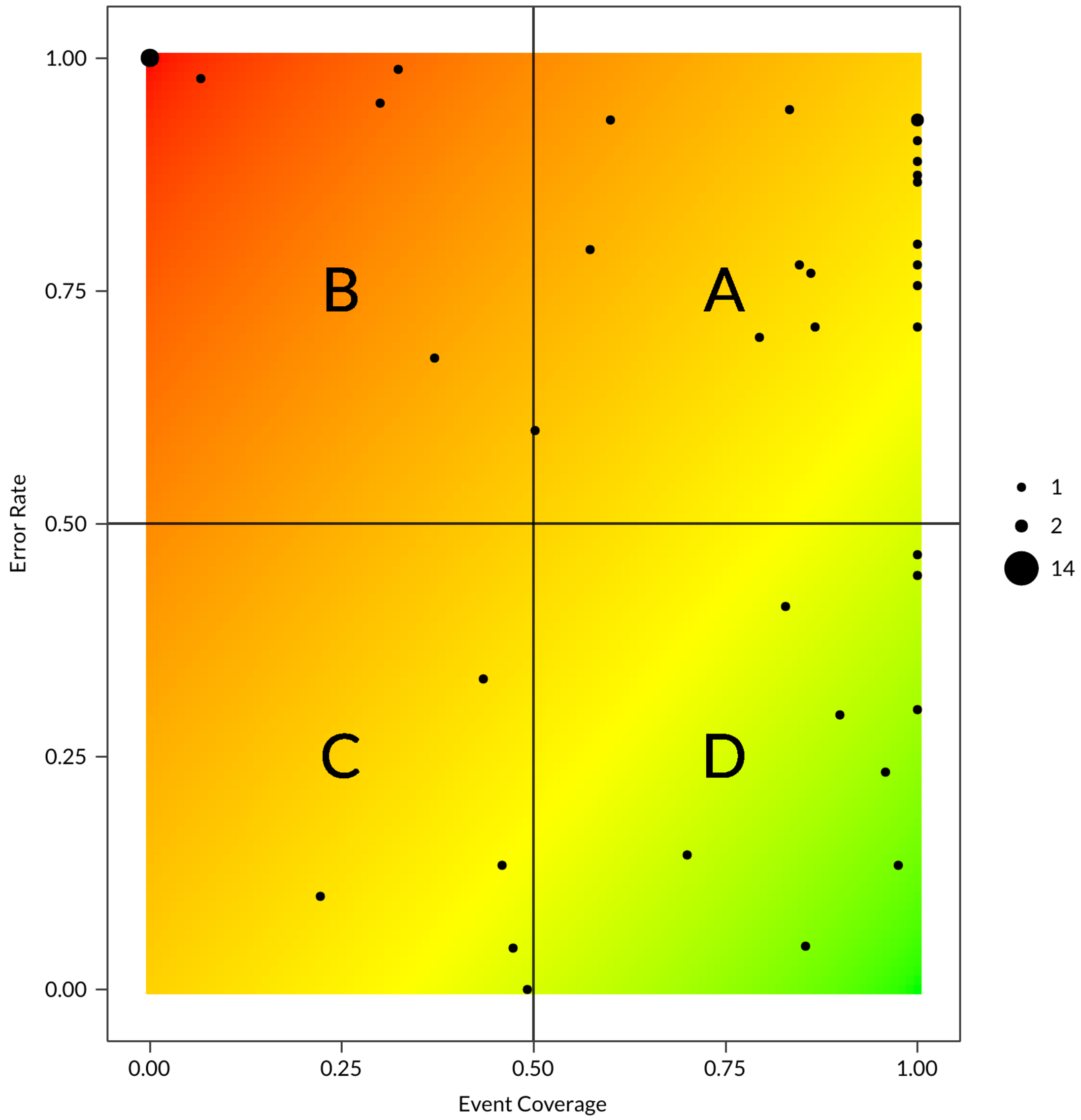


Figure 5.4: Four-field graph of the event coverage, error rate, and indicator values with the value points calculated from the data.

Chapter 6

Evaluation

6.1 Quality of Traffic Information in Finland

Incident timeliness (start) criterion has a basic one-star quality level. The 95 % quantile value of the criteria indicates that the effort to achieve the three-star advanced level of quality requires substantial changes in the incident reporting process. The criterion value distributions show logical skewness to the right as slower detection time is more unlikely to occur. The elevated frequencies around 30-minute criterion value form an exception. This can be the result of the 30-minute maximum response time for incident impulse queue for the HäTi operator as large share of the reports are done just before the maximum response time mark.

Latency criterion depends mainly on the time in which the Road Traffic Department operator opens the incident input dialog from the incident impulse queue. The 95 % quantile of the criterion indicates that the reporting process is nowhere close to the lowest one-star quality level, which can only be achieved with faster operation of the impulse queue. This means that either the application used to operate queues should be operated faster or at least the preliminary announcement should be made automatically. The frequency distribution of the latency criterion is strongly skewed to the right and the frequencies drop at 30-minute because of the same required maximum queue operation time. While the largest frequency peak is between 1–10 minutes, some values get as large as 50 minutes. These outliers can be the result of 60 minute tolerances used when joining FERCA and Digitraffic cases. As stated in chapter 5, 20 % of incident cases have no preliminary announcements. From these cases, over 75 % have the latency value of over 10 minutes which is the maximum required time for the lowest quality level. Therefore, the lowest quality level could be achieved with more consistent preliminary incident reporting.

Location error across all incidents reports is fairly low – the corresponding criteria quality level is two stars. The value is calculated in a different way than expected in the criteria; it is distance between two end points on the incident road segment and the corresponding ground truth segment. This means that the location error is greatly dependent on the network geometry and measurement link lengths. Therefore, location error is not a trustworthy measure of quality as long as incident location is relaxed to long segments. The segmentation of the criterion value calculation can also be seen in the histogram where a large share of cases accumulate to the low side of the graph. The phenomenon is a result of a consistent placement of the detected and Digitraffic incidents on the same HERE measurement link.

Error rate criteria variants have mixed results: while the EIP defined criterion average value is 13 %, the QRTTI variant is as high as 57 %. On the other hand the respective values with the non-timeout incidents are 7 % and 42 %. This suggests that the incidents with the 90 minute timeout value in cases where there is no message indicating the end of the situation have a negative influence on the quality of the incident reporting. This is because the criteria value is a monotonically decreasing subject to the message area size. One way to address this issue is to define the timeout more intelligently and not to use a constant value. The timeout could be estimated as a function, which has parameters for different hours in a day, the amount of traffic, the overall capacity and the type of the incident by studying the non-timeout incident reports in the past. The distribution histogram of the criterion value is weighted towards the high-end side of the value range. This indicates that most of the Digitraffic incidents overlap or miss the ground truth area by a wide margin in both spatial and temporal axis.

The EIP defined variant of the event coverage average value is 87 % and the QRTTI variant is 62 %. The respective values with the non-timeout incidents are 93 % and 57 %. Both values remain relatively same in both variants which suggests that the timeout usage has little effect on the event coverage. The better QRTTI variant value would need very precise incident areas. Ultimately, the perfect QRTTI event coverage and error rate scores can only be achieved by having the exactly duplicate ground truth and incident message areas. If the previously proposed intelligent timeout value works well, it has only a positive effect on the error rate and no negative effects on the event coverage. The frequency distribution of the QRTTI variant is roughly uniform with an exception at the 100 % value. While also considering the accumulation of high error rate values, the incident reports have tendency to have overly large ranges for start and end times and locations. The reason for low scores for the QRTTI variant of the event coverage criterion is that the FTA does not intend to take the whole incident impact zone into account in the incident reporting instead only reporting the location of the incident. For better criteria scores in the future, the impact zone should be taken into account. Additionally, incidents that were both apparent in the ground truth and FERCA data but not in Digitraffic should be investigated.

The quality of the all listed criteria mainly depends on two factors. The human factor consists of the speed and precision in which the emergency operator registers the incident in the FERCA systems, the emergency personnel provides the additional details of the incident to the road traffic department, and the FTA operator maintains the impulse queue in the HäTi system. The impact of this factor to the quality can be optimised by training and by maintaining the vigilance of the personnel with meaningful work shift lengths and the sufficient number of breaks. Additionally, sufficient work resources should be maintained even in busier situations, such as large events and in a bad weather. The second, systematic factor includes the operating speed and reliability of the computer systems involved in the incident reporting process. This factor can be refined by actively updating

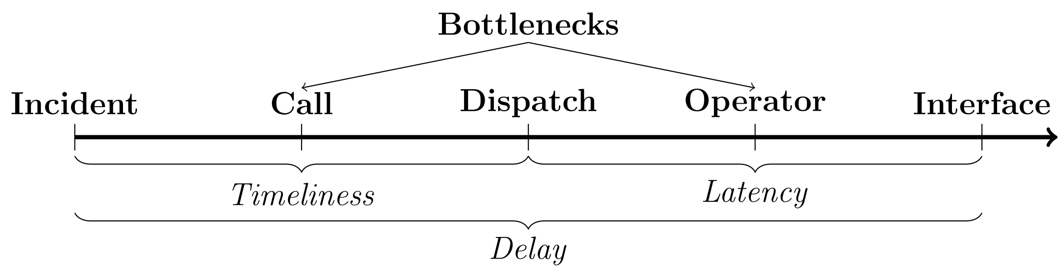


Figure 6.1: Possible bottlenecks in the incident information process.

the usability and the speed of the said systems.

The reaction time in which a bystander makes an emergency call to the emergency centre and the incident impulse queue operating speed by the HäTi operator form the potential bottlenecks in the incident information process. These bottlenecks can be eliminated by automating the respective tasks. The first detection of the incident should be based on a detection system which rapidly detects possible traffic incidents by analysing the speed of the traffic in real time. The incident detection algorithm used in this thesis could be used as a start point for the process that could be used in the future. The second part in the automation is to address the delay of the incident impulse queue operation in the HäTi system. The preliminary incident announcement could be sent automatically provided that the FERCA operator has sufficient information for referencing the location of the incident more accurately. This could be made possible in the future when automatic, GPS-utilising emergency call systems in road vehicles are a standard equipment.

Overall, the level of traffic incident reporting quality in Finland fulfills the minimum, basic requirements stated by the EIP and the ITS directive. While event coverage and error rate values are very challenging, other criteria show promising results. Therefore, it is safe to assume that FTA achieves at least basic one-star quality from the overall viewpoint.

6.2 EIP Quality Criteria

Traffic incident reporting in Finland has no strict concept of incident validation. Basically, an incident is deemed validated after the distress call and emergency personnel dispatch from the FERCA. Therefore, the two-star level of quality in the Timeliness criterion is not compatible for traffic incident reporting quality assessment in Finland. Generally, the 'Validation (if necessary)' clause in the timeliness criterion definitions is not coherent, as in the case of missing validation no two-star level of quality is available. This thesis suggests that an alternative two-star level of quality criterion is defined for non-validating reporters. This level should be defined similarly to the three-star definition of detection and validation except the timeliness value limit should be higher.

Location accuracy in the criteria is defined as the euclidean distance between the reported and true location of the traffic incident. In Finland the majority of the incident report locations are referenced as a road section between two TMC points, or more generally, intersections. Location error can be calculated from the distance of the end points in the report and the ground truth. This value is not compatible with the level of quality definitions.

As stated in the study of the recommended quality criteria in section 3.1, incident reporting quality measures should not only consist of external, result-centric requirements but also structural quality factors. One important structural quality concern in open data is the ease of use of the data retrieval interface. As a whole, DATEX II interface and the format used by Digitraffic proved to be very complex to process. While the said format is very versatile and includes many different fields and options for traffic status reporting, much of the retrieved data was redundant and sometimes erroneous in certain fields.

6.3 Evaluation of the Methods

The number of detected incidents is very high, near $95,000$. Most of the incidents are false alarms, such as inexplicable high falls in traffic speeds because of a road block, unexpected braking, or traffic lights. Even poor weather conditions are suspected to be a catalyst for false detections.

Despite the high number of detected incidents, the accuracy of the incident detection algorithm proved to be relatively low. One significant limitation of the incident detection algorithm is that it only performs adequately in high traffic road sections. Only 50 % of the incidents listed in Digitraffic on the study period and location delimitations were detected. Most of the missing incident detections resulted because of insufficient amount of traffic, or HERE probes, at the time of the incident. This was particularly apparent in low-traffic moments in the day. In high-traffic environments the detection algorithm performed with acceptable performance, although in extreme cases it had tendency for false detections. One possible way to enhance the detection algorithm is to introduce additional data to the process, such as speed and flow data from other floating car data sources. The information about the occupancy of the roads could also greatly improve the detection performance. The only source providing this information is the LAM system which is too sparse for a effective use.

As the FTA incident reports are mapped to the HERE network with a lower accuracy, the length of the HERE links play a significant role in location accuracy as well as QRTTI-related criteria. If the corresponding HERE measurement links at the vicinity of the traffic incidents are very long, this results in higher volatility of the location accuracy, error rate, and event coverage depending on the compatibility of the ground truth and the incident message area. In areas where the HERE network granularity is high, this problem is greatly reduced. Additionally, the uneven granularity in the network results in incompatible quality values across different areas. This problem could be solved by either using a ground truth that provides incident locations in same location reference as the incident reports or by redefining the quality criteria to better fit the relaxed location references.

Both error rate and event coverage criteria are calculated using the whole segment-of-impact for traffic speeds as a ground truth instead of only using the actual measurement link where the accident has occured. This greatly reduces both mentioned criteria values. Currently, FTA does not intend to include traffic impacts in the incident reports while the QRTTI variant of error rate and event coverage criteria assess the quality from the viewpoint of an end user. After all, the road user is usually not only interested in the location of the traffic incident, but also in the possible traffic congestion caused by the incident.

The speed of the detection algorithm is very slow. The overall process of detection, combining, aggregating and joining lasts up to 3 hours per configuration. One possible way to optimise the performance of the algorithm is to better utilise parallel processing. Many parts of the code are sequential and single threaded which can be improved by refactoring the code to use all 64 cores available in the hardware used thus greatly reducing the operation time of the detection algorithm.

Chapter 7

Discussion

7.1 The Value of the Developed Methods to Stakeholders

Methods introduced in this thesis are likely to provide at least implicit value to several stakeholders in both public and private sectors as well as end users.

Because the research objective for this thesis was to implement quality assessment and assurance methods for the FTA incident reporting, it is considered as the main beneficiary from the associated findings. FTA can optimise their incident information value chain by considering changes proposed in chapter 6. They can also augment or change their quality assurance routines with principles introduced in chapter 4.

The EIP member states get valuable information about the usefulness of the chosen ground truth as incident detected from HERE speed data apply to all EU countries. As a result, the same source ground truth can also be used elsewhere. This adheres to the harmonisation objective stated in the ITS directive.

The RTTI criteria and the associated quality assessment and assurance methods could also be utilised in the private sector. Commercial real-time traffic information services could utilise their own traffic speed data for incident detection and use the result data as a ground truth for quality assessment. Quality assurance methods could also apply to commercial parties if they provide additional value for existing processes. Detected incidents can also augment the existing traffic incidents for map and navigation service providers.

In the near future, road users benefit mostly from the possible traffic information quality improvements provided by the FTA followed by corrective actions provided in chapter 6. For example, non-recurring congestions can be better avoided with more timely and precise incident information.

7.2 Ways to Improve Quality Assessment

The selection of ground truth for traffic incident information quality assessment proved critical for the success of the study. There are several commercial real-time traffic information providers that could be used instead. One alternative could be seen equal to the HERE platform. TomTom also has wide coverage of countries and large user base for a worldwide incident detection algorithm use. The even more popular alternative Google Maps could be the best choice as it has a very large user base. On the other hand, the interface of this consumer level provider is very closed and only provide pre-calculated congestion indices instead of raw data.

7.3 Application to Real-Time Incident Detection

While the quality assessment methods in this thesis are designed as per EIP+ requirements, some of the findings may provide an assisting functionality for other applications. The detection algorithm can be utilised for real-time incident detection after a few modifications. The time window approach is not optimal for detecting incidents in real time. This can be addressed by studying the most recent speed value. If the value drops below estimated confidence intervals, the associated road segment is deemed as a suspect for traffic incident. After a sufficient amount of low speed values the incident is deemed real and necessary actions are performed, such as sending an automatic preliminary incident report. The detected incidents could provide base for a more timely preliminary incident reporting. While the sensitivity of the incident detection is quite high, this could be addressed by altering the detection sensitivity parameters.

7.4 Technological Changes and the Future of Quality Assessment

The traffic surveillance and personal equipment technology is constantly evolving. This provides new possibilities for traffic information quality assessment studies. Starting from year 2018, all new road vehicles are equipped with an automatic eCall system [16] that makes a phone call to the emergency centre along with the vehicle coordinates in case of an accident. While the renewal of the nationwide fleet of cars is slow in Finland, the future promises more efficient and reliable incident detection. The car fleet renewal problem could be resolved by inexpensive retrofitting of eCall systems to old road vehicles.

In the near future, FTA is likely to replenish their sources for incident data collection. One possible way to gather additional information about traffic incidents is to use crowd sourcing. This can be done either actively by providing a mobile application where road user can submit witnessed incidents or by filtering and manipulating data from popular social media platforms, such as Twitter and Facebook. A new traffic flow measurement process called SUJUVA [26] is also being developed. In SUJUVA, the traffic flow, speed, and occupancy are measured using real-time data derived from mobile phone cell locations. SUJUVA could provide large samples for travel time estimation and incident detection, given that cell location measurements are sufficiently accurate.

Finland is considering the introduction of road tolls and congestion pricing particularly in the capital region. While the main objectives for these systems are to decrease traffic in congestion sensitive road sections and provide funding for transport investment and operations, they can also provide traffic volume data for analytic purposes. The value of this data will depend on the chosen technological solution.

It is possible that the connectivity in road vehicles develop into a state where no central database is required. When a traffic incident or an other type of disturbance occurs, the information about the incident is sent by an eCall-like system not only to the emergency centre but also to nearby road vehicles. The information could also propagate to other vehicles via a car-to-car network. This could result in a more passive, supervising role for the government in the traffic information provision. As a result, traffic information quality assessment and assurance have to adapt to decentralised data collection.

Several automobile and consumer electronics companies are developing autonomous vehicles which could enable driverless commuting to road users. Autonomous operation requires various sensory inputs from the environment as well as information about the state of the nearby traffic.

Due to future developments in the traffic incident reporting context, the incident information value chain is subject to change. The content side of the value chain is radically simplified by only containing automatic distress calls from eCall systems and the provision of the incident information to nearby road users. The content segment will also be reduced to more simple process where the road vehicles have a responsibility of information processing and presentation.

Chapter 8

Conclusions

This thesis studied the quality of the traffic incident information provided by the Finnish Transport Agency. The quality assessment was performed with a set of criteria defined in the European ITS Platform (EIP) project invoked by the ITS directive adopted by the European Parliament and the Council.

The main objective of this thesis was to implement a set of quality assessment methods for real-time incident information based on the EIP criteria. All criteria were divided into four levels of quality that have explicit required value levels.

Another task in this thesis was to construct a basic quality assurance process for continued quality monitoring. This process will provide a simple way to analyse the overall quality of the incident information and provide additional information to determine the source of a possible problem.

The studied incident information was retrieved from the FTA Digitraffic interface. For primary ground truth, traffic incidents were detected from HERE speed data. Additionally, emergency missions recorded by the Finnish Emergency Response Centre were acquired. The incidents detected from HERE form the reference in which all the data is tested against. The FERCA data represents the first moment in time when the incident is detected.

The listed criteria, used data and defined methods are listed in table 8.1. For timeliness, the time difference of timestamps of the emergency mission times from Finnish Emergency Response Centre data and detected incident start times from HERE speed data were calculated. The latency criterion was calculated as a time difference between the Digitraffic and FERCA timestamps. The location accuracy criterion has been divided into two different criteria values. In first and second levels of quality the presence of the incident location in the correct link between intersections detected from HERE was studied. In the third and fourth levels the location accuracy was defined as a mean distance in kilometres between the start and end points of the HERE and Digitraffic incidents. In error rate and event coverage, Digitraffic and HERE data were aligned to a two-dimensional spatio-temporal space by referenced and detected road segments and the duration of the incident. The error rate value was calculated as a proportion of the reported incident area not covering the ground truth. The event coverage value was calculated as proportion of the detected incident area that is correctly reported.

The criteria calculations were first studied as mean values to better understand the overall value of the incident information published by the FTA. The final

Table 8.1: Quality criteria, used data, and methods.

Criterion	Used Data	Method
Timeliness	HERE, FERCA	Time difference (FERCA - HERE)
Latency	FERCA, Digitraffic	Time difference (Digitraffic - FERCA)
Link Between Intersections	HERE, Digitraffic	Reported location in detected most downstream incident link
Location Accuracy	HERE, Digitraffic	Mean distance between the start and end points
Error Rate	HERE, Digitraffic	The proportion of the reported incident area not covering the ground truth area
Event Coverage	HERE, Digitraffic	The proportion of the ground truth area covered by the incident report

results along with the associated quality levels were calculated from table 1.1. In timeliness, latency and location accuracy criteria, the defined levels must be met by 95 % of the calculated values. In other criteria the mean values were used directly. The criteria quality levels were listed in table 5.1. The highest achieved level of quality is the two-star level which was achieved with the location accuracy criterion. Basic one-star quality level was given to EIP defined error rate and event coverage as well as timeliness criteria. The latency criterion did not achieve the lowest quality level.

The overall level of the traffic incident information quality is relatively low. There are several reasons for these results. First of all, the detection algorithm did not perform with acceptable precision because HERE speed data alone was not sensitive enough for traffic incidents. The outcome of this was that the sample size for the assessment was not optimal resulting in biased calculations which could explain low quality ratings. More information about the traffic occupancy and quantity could have helped with detection. Second, all criteria definitions were not suitable for Finnish environment. For example, the higher quality levels of the location criterion required that spatial calculations should have been performed with coordinate points instead of road sections. This resulted in values that depended on the length of the HERE network measurement links. Third, the comparatively high incident timeout value of 90 minutes had great impact on QRTTI variants of error rate and event coverage criteria. While the usage of timeout caused better event coverage values, higher error rate values resulted in weaker overall quality. Last, 20 % of the incident cases had missing preliminary reports, which had considerable, deteriorating effect in latency criterion values.

Despite drawbacks mentioned in this thesis, the detection algorithm has a lot of potential. The formation of the ground truth data is very difficult in Finland, as potential data sources are sparse. Therefore, incidents detected from traffic data form one of few possible candidates for the ground truth. As a result, the incident detection method should be developed further by studying underlying principles more closely and considering additional types of data sources for better results.

If the accuracy of the detection algorithm can be improved, it could also hold value in other traffic-specific areas. With modifications, the algorithm could detect traffic incidents in real time and thus serve as a data source for preliminary traffic incident reports. This requires that the detection sensitivity should be lowered to more realistic level to avoid unnecessary alerts.

References

[1] Breakout detection in the wild. https://blog.twitter.com/2014/breakout-detection-in-the-wild. Accessed: March 24, 2015.

[2] DATEX II. http://www.datex2.eu/. Accessed: July 3, 2015.

[3] Digitraffic. http://www.infotripla.fi/digitraffic/doku.php?id=start_en. Accessed: Mar 24, 2015.

[4] Finnish emergency response centre. http://www.112.fi/other_languages. Accessed: Mar 25, 2015.

[5] HERE Platform. https://developer.here.com/. Accessed: Mar 25, 2015.

[6] HäTi interview with Jani Laiho from FTA. Personal communication. Interview date: Dec 12, 2014.

[7] Liikennetilanne. http://liikennetilanne.liikennevirasto.fi/. In Finnish. Accessed: Jan 9, 2015.

[8] Liikenteen automaattinen mittausjärjestelmä. http://alk.tiehallinto.fi/alk/info/lilakuvaus.htm. In Finnish. Accessed: Mar 3, 2015.

[9] Local principal curve methods. http://cran.r-project.org/web/packages/LPCM/index.html. Accessed: March 24, 2015.

[10] Nonparametric multiple change-point analysis of multivariate data. http://cran.r-project.org/web/packages/ecp/index.html. Accessed: March 24, 2015.

[11] QSRTI interview with Teun Hendriks from TNO. Personal communication. Interview date: Jan 13, 2015.

[12] The R project for statistical computing. http://www.r-project.org/. Accessed: Jan 4, 2015.

[13] SQLite interface for R. http://cran.r-project.org/web/packages/RSQLite/index.html. Accessed: March 24, 2015.

[14] Support functions and datasets for venables and ripley's MASS. http://cran.r-project.org/web/packages/MASS/index.html. Accessed: March 24, 2015.

[15] Tools for parsing and generating XML within R and S-Plus. `http://cran.r-project.org/web/packages/XML/index.html`. Accessed: March 24, 2015.

[16] Type-approval requirements for the deployment of the eCall in-vehicle system based on the 112 service. `http://www.europarl.europa.eu/oeil/popups/ficheprocedure.do?lang=en&reference=2013/0165%28COD%29#tab-0`. Accessed: Aug 31, 2015.

[17] V-Traffic – Koko Suomen reaaliaikainen liikennetieto. `http://v-traffic.fi/`. In Finnish. Accessed: Jan 9, 2015.

[18] Directive 2010/40/EU of the European Parliament and of the Council on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport. Official Journal of the European Communities, July 2010.

[19] Ahmed, S., and Cook, A. Application of time-series analysis techniques to freeway incident detection. *Transportation Research Record*, 841 (1982).

[20] Bogenberger, K., and Weikl, S. Quality management methods for real-time traffic information. *Procedia Social and Behavioral Sciences 54* (2012), 936–945.

[21] Cook, A., and Cleveland, D. Detection of freeway capacity-reducing incidents by traffic-stream measurements. *Transportation Research Record 495* (1974), 1–11.

[22] European ITS Platform. *Proposal for Quality Assessment of ITS Services on the TERN* (2015).

[23] Everitt, B. S., Landau, S., Leese, M., and Stahl, D. *Cluster Analysis*, 5th ed. John Wiley & Sons, Chichester, United Kingdom, 2011.

[24] Hamilton, J. D. *Time Series Analysis*, vol. 2. Princeton University Press, Princeton, NJ, USA, 1994.

[25] Hendriks, T., and van der Vliet-Hameeteman, C. QSRTI: a method for quality assessment of event based traffic information with a focus on Safety Related Traffic Information. Tech. rep., TNO, 2014.

[26] Innamaa, S., and Hätälä, E. Floating mobile data pilot in the Helsinki Metropolitan Area. Tech. rep., VTT, 2012.

[27] ISO 14819-2:2013. *Part 2: Event and information codes for Radio Data System – Traffic Message Channel (RDS-TMC) using ALERT-C*. ISO, Geneva, Switzerland.

[28] Kan, S. H. *Metrics and Models in Software Quality Engineering*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2002.

[29] Montgomery, D. C. *Introduction to Statistical Quality Control*, 6th ed. John Wiley & Sons, Chichester, United Kingdom, 2009.

[30] Razali, N. M., and Wah, Y. B. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. *Journal of Statistical Modeling and Analytics 2*, 1 (2011), 21–33.

[31] Steenbruggen, J., Kusters, M., and Broekhuizen, G. Best practice in european traffic incident management. *Procedia-Social and Behavioral Sciences 48* (2012), 297–310.

[32] Turner, S. Defining and measuring traffic data quality: White paper on recommended approaches. *Transportation Research Record: Journal of the Transportation Research Board 1870*, 1 (2004), 62–69.

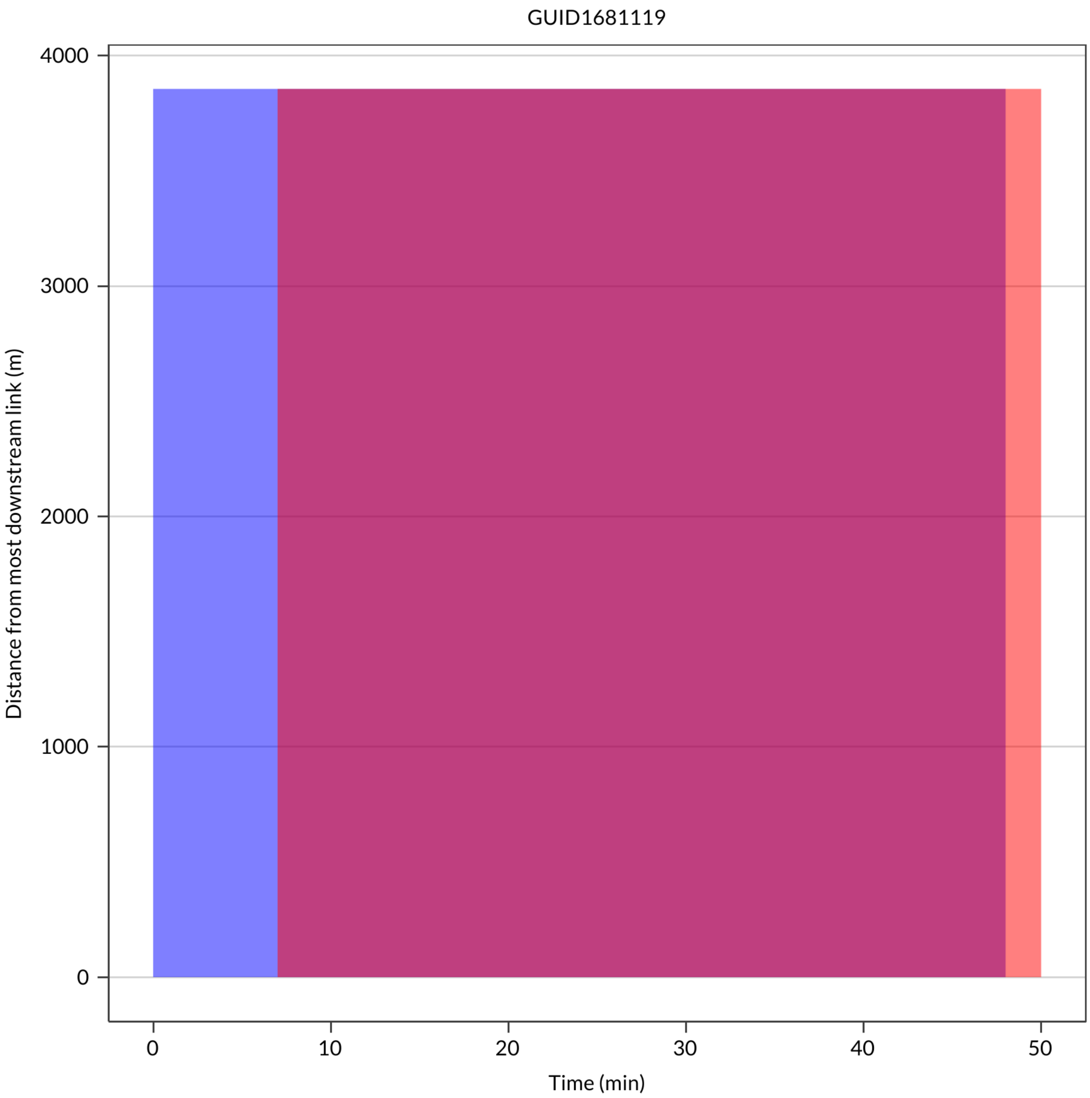
Appendix A

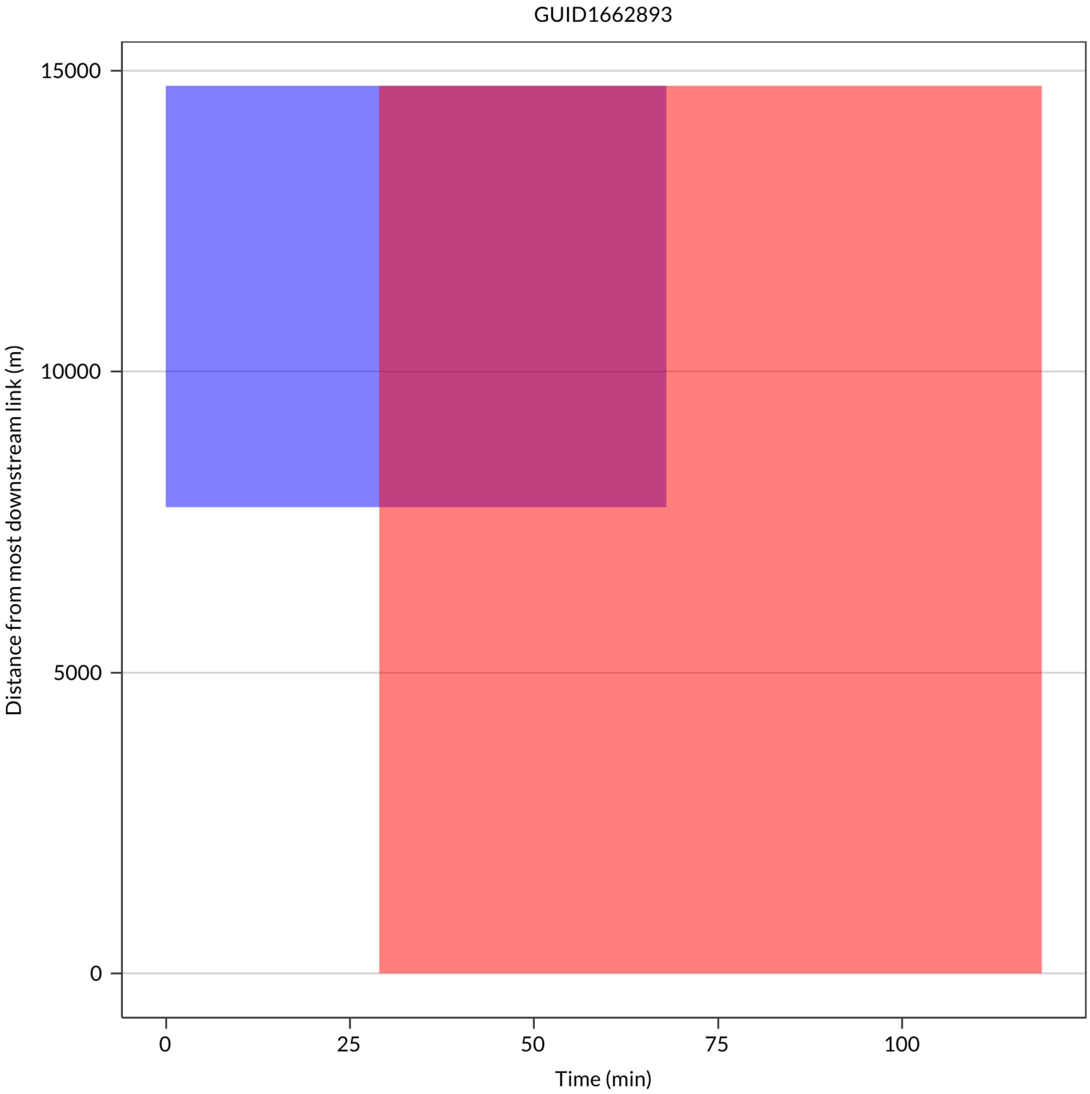
DATEX2/AlertC Example

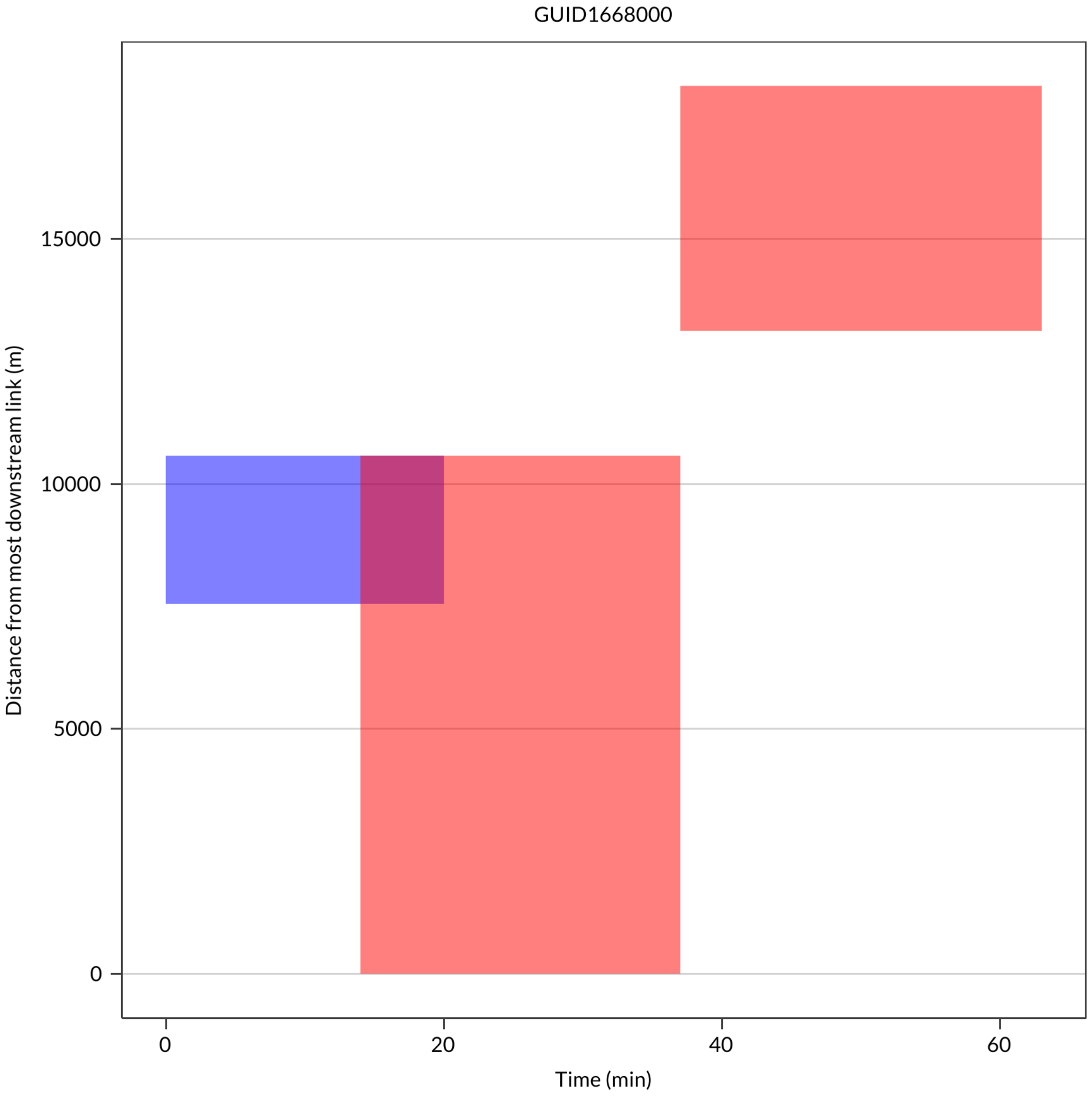

```
<situation id="GUID1631012" version="1">
  <headerInformation>
    <confidentiality>restrictedToAuthoritiesTrafficOperatorsAndPublishers</confidentiality>
    <informationStatus>real</informationStatus>
  </headerInformation>
  <situationRecord id="GUID463496901" version="1" type="Accident">
    <situationRecordCreationTime>2014-10-01T14:29:34+03:00</situationRecordCreationTime>
    <situationRecordVersionTime>2014-10-01T14:29:34+03:00</situationRecordVersionTime>
    <situationRecordFirstSupplierVersionTime>
        2014-10-01T14:29:34+03:00
    </situationRecordFirstSupplierVersionTime>
    <probabilityOfOccurrence>certain</probabilityOfOccurrence>
    <validity>
      <validityStatus>definedByValidityTimeSpec</validityStatus>
      <validityTimeSpecification>
        <overallStartTime>2014-10-01T11:46:13</overallStartTime>
        <overallEndTime>2014-10-01T14:29:59+03:00</overallEndTime>
      </validityTimeSpecification>
    </validity>
    <generalPublicComment>
      <comment>
        <values>
            <value lang="fi">
                Tie 8, Rauma. Onnettomuuspaikan raivaustyö on ohi.Tie 8 välillä Turku - Pori.
                Paikka: Rauma. Noin 47 km ennen paikkaa Pori.Tarkempi paikka: Välillä
                Rauma - Rauman eteläinen liittymä. Onnettomuuspaikan raivaustyö on ohi.
                Tapahtuma ohi Kesto: 01.10.2014 klo 11:46 - 14:29
                http://liikennetilanne.liikennevirasto.fiLiikenneviraston
                tieliikennekeskus TurkuPuh 0206373329Faksi 0206373711Sähköposti
                turku.liikennekeskus@liikennevirasto.fi
            </value>
        </values>
      </comment>
    </generalPublicComment>
    <groupOfLocations type="Linear">
      <alertCLinear type="AlertCMethod2Linear">
        <alertCLocationCountryCode>6</alertCLocationCountryCode>
        <alertCLocationTableNumber>17</alertCLocationTableNumber>
        <alertCLocationTableVersion>1.11.03</alertCLocationTableVersion>
        <alertCDirection>
          <alertCDirectionCoded>Both</alertCDirectionCoded>
        </alertCDirection>
        <alertCMethod2PrimaryPointLocation>
          <alertCLocation>
            <specificLocation>1621</specificLocation>
          </alertCLocation>
        </alertCMethod2PrimaryPointLocation>
        <alertCMethod2SecondaryPointLocation>
          <alertCLocation>
            <specificLocation>1622</specificLocation>
          </alertCLocation>
        </alertCMethod2SecondaryPointLocation>
      </alertCLinear>
    </groupOfLocations>
    <accidentType>other</accidentType>
  </situationRecord>
</situation>
```

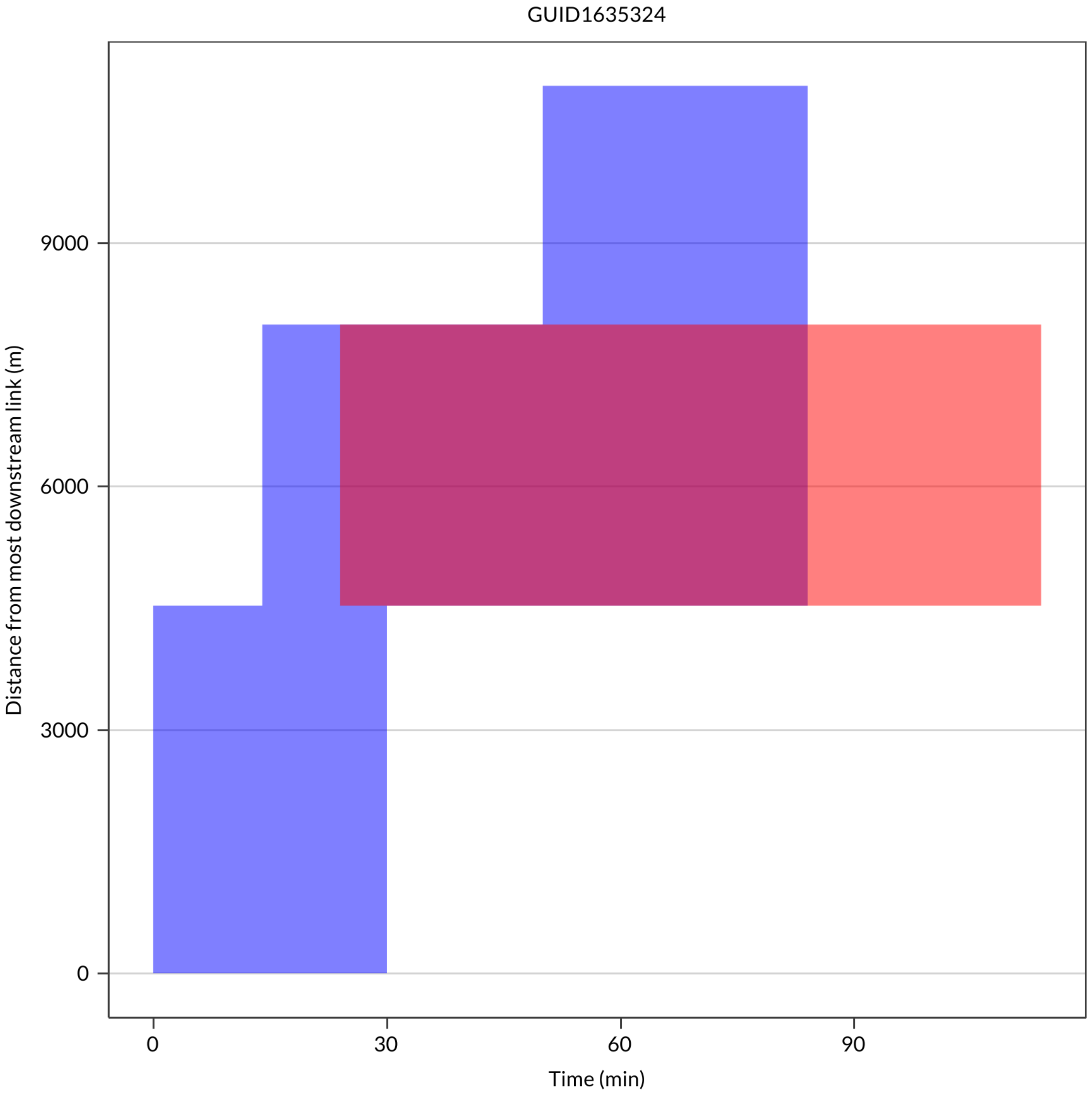
Appendix B

Example QRTTI Graphs



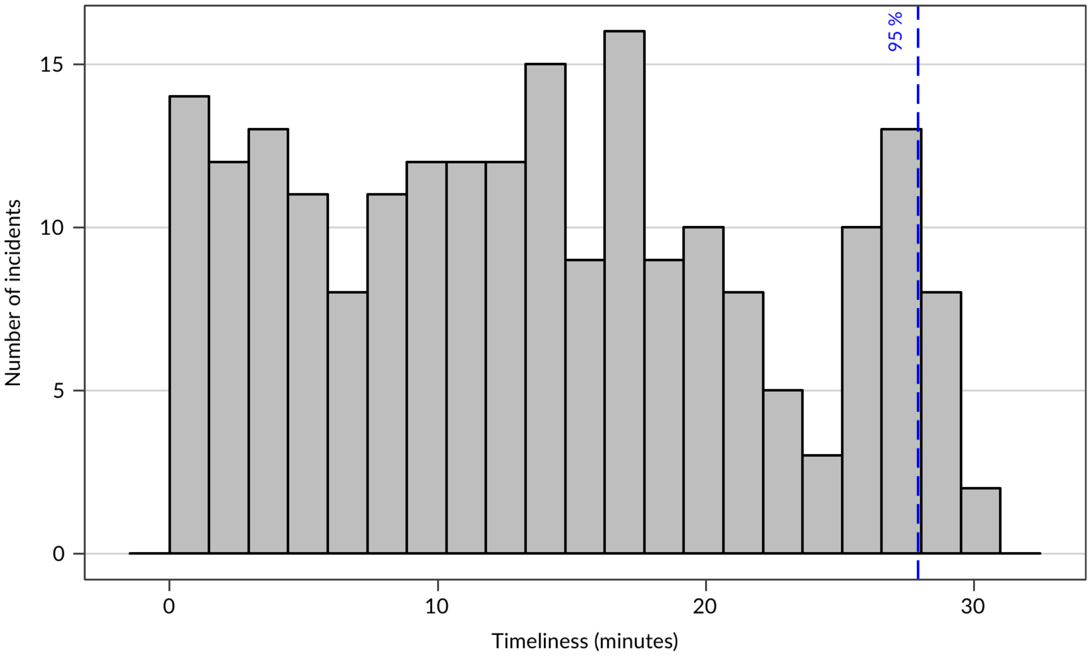

GUID1662893
15000
10000
5000
0
Distance from most downstream link (m)
0
25
50
75
100
Time (min)

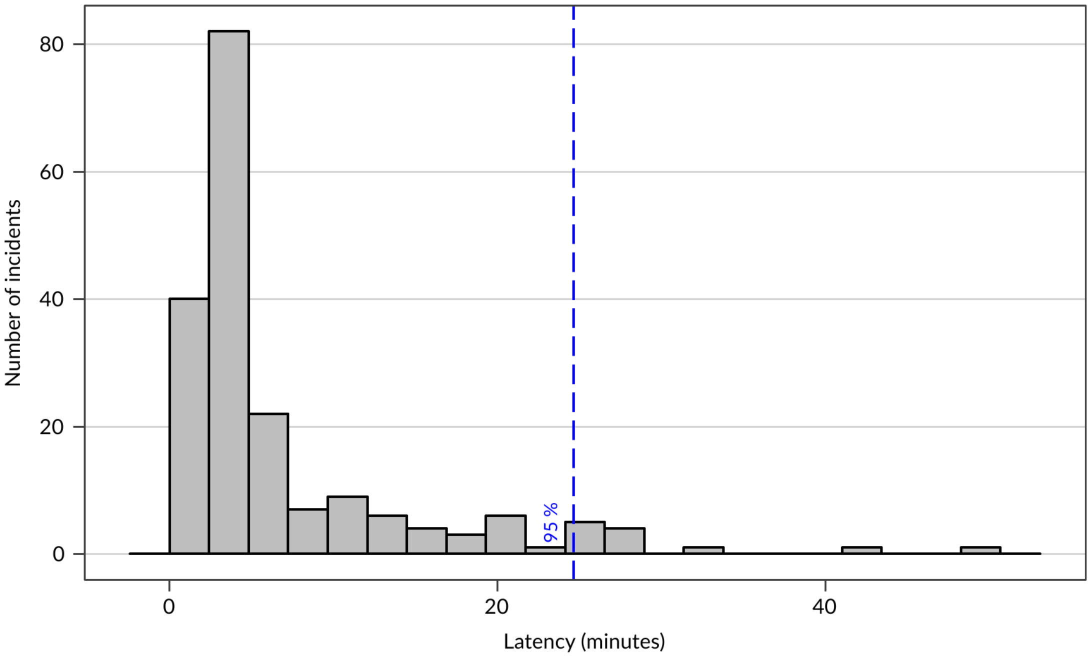

GUID1668000
Distance from most downstream link (m)
15000
10000
5000
0
0
20
40
60
Time (min)

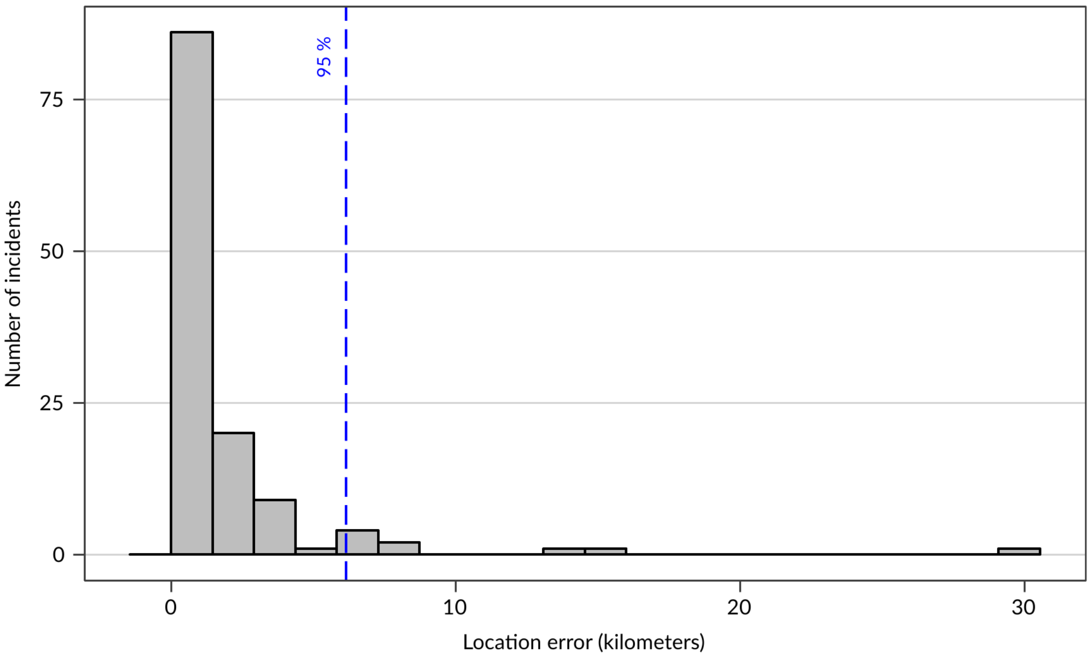

GUID1635324
9000
6000
3000
0
Distance from most downstream link (m)
0
30
60
90
Time (min)

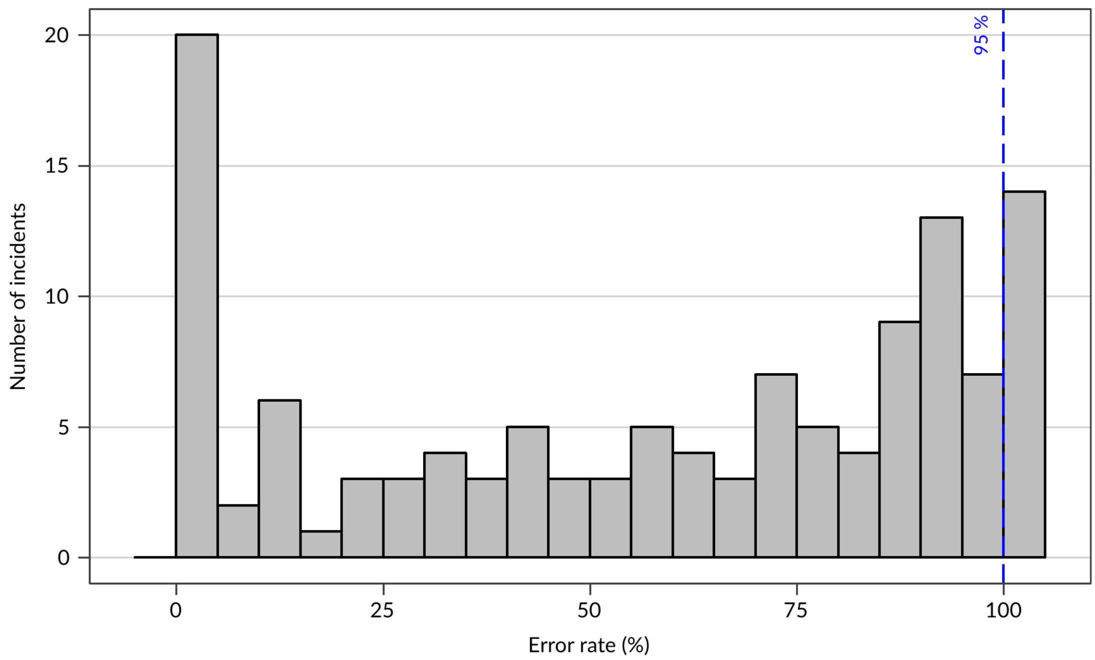
Appendix C

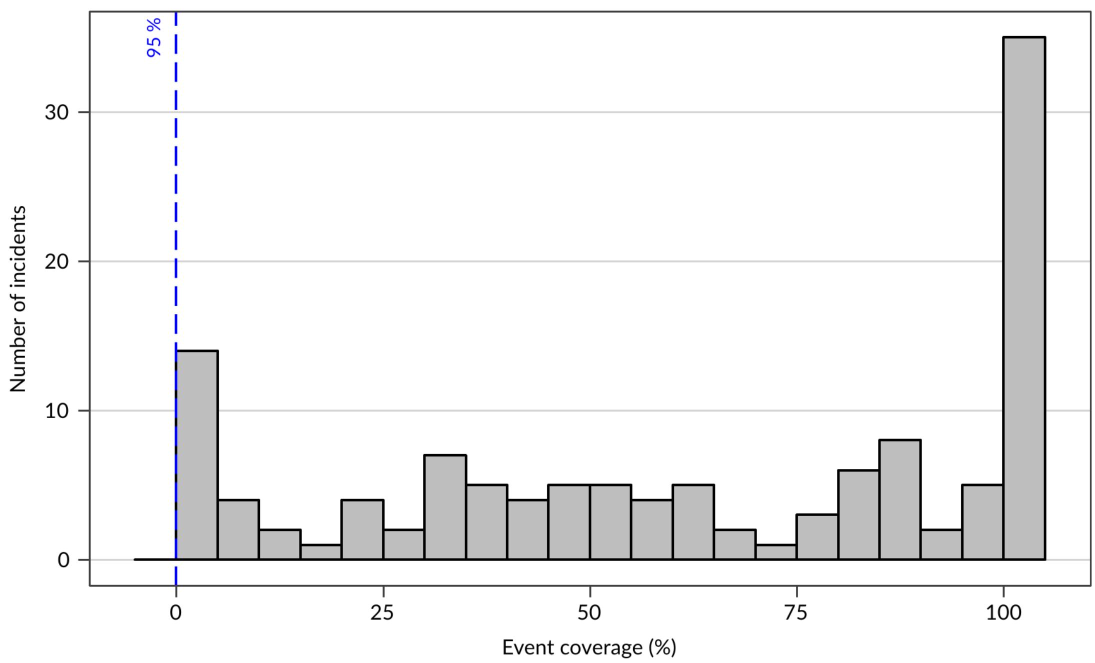
Criteria Value Histograms




Number of incidents
80
60
40
20
0
95 %
0
20
40
Latency (minutes)


Number of incidents
75
50
25
0
95 %
0
10
20
30
Location error (kilometers)


95 %
Number of incidents
0
5
10
15
20
0
25
50
75
100
Error rate (%)


95 %
Number of incidents
0
10
20
30
0
25
50
75
100
Event coverage (%)